ASHGATE

Modeling Applications in the Airline Industry

Ahmed Abdelghany and Khaled Abdelghany

ASHGATE e-BOOK

MODELING APPLICATIONS IN THE AIRLINE INDUSTRY

This page has been left blank intentionally

Modeling Applications in the Airline Industry

AHMED ABDELGHANY
Embry-Riddle Aeronautical University, USA

&

KHALED ABDELGHANY
Southern Methodist University, USA

ASHGATE

Published by

Ashgate Publishing Limited
Wey Court East
Union Road
Farnham
Surrey, GU9 7PT
England

Ashgate Publishing Company
Suite 420
101 Cherry Street
Burlington
VT 05401-4405
USA

www.ashgate.com

British Library Cataloguing in Publication Data

Abdelghany, Ahmed F.
Modeling applications in the airline industry.
1. Airlines--Management--Simulation methods.
I. Title II. Abdelghany, Khaled.
387.7'068-dc22

ISBN: 978-0-7546-7874-8 (hbk)
978-0-7546-9725-1 (ebk) I

Library of Congress Cataloging-in-Publication Data

Abdelghany, Ahmed F.
Modeling applications in the airline industry / by Ahmed Abdelghany and Khaled Abdelghany.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-7546-7874-8 (hardback) -- ISBN 978-0-7546-9725-1 (ebook)
1. Airlines--Management. 2. Aeronautics, Commercial--Computer simulation. 3. Aeronautics, Commercial--Planning. 4. Operations research. I. Abdelghany, Khaled. II. Title.
HE9780.A228 2009
387.7068'4--dc22

2009017554

FSC
Mixed Sources
Product group from well-managed forests and other controlled sources
www.fsc.org Cert no. SA-COC-1565
© 1996 Forest Stewardship Council

Printed and bound in Great Britain by
MPG Books Group, UK

Contents

List of Figures *vii*
List of Tables *xi*

1 Introduction to Airline Management 1

SECTION I DEMAND MODELING AND FORECASTING

2 Modeling the Choice of Travel Options 21

3 Passenger Demand Modeling and Forecasting 39

SECTION II SCHEDULING OF RESOURCES

4 Fleet Assignment 53

5 Aircraft Routing 79

6 Crew Planning 89

7 Gate Assignment 109

8 Baggage Handling 119

9 Flight Planning and Fuel Management 129

SECTION III REVENUE MANAGEMENT

10 Introduction to Revenue Management 141

11 Demand Forecasting for Revenue Management 155

12 No-show Rate and Overbooking 173

13 Seat Inventory Control for Flight-based Revenue Management Systems 181

14 Seat Inventory Control for Network-based Revenue Management Systems 189

15	Ticket Distribution	205
16	Sales Contracts	213
17	Code-share Agreements	221
SECTION IV	**IRREGULAR OPERATIONS MANAGEMENT**	
18	Ground Delay Programs and Collaborative Decision Making	229
19	Impact of Disruptions on Air Carrier Schedule	243
20	Airline Schedule Recovery	253
Index		*275*

List of Figures

Figure 1.1	The different players in the air transportation industry	3
Figure 1.2	Decision levels of airline management	5
Figure 1.3	Changes in the main market characteristics due to the 9/11 terrorist attack in the domestic US market	6
Figure 1.4	Processes considered in the planning and operations phases of the airlines	9
Figure 1.5	Example of a hub-and-spoke network structure	10
Figure 1.6	Example of time bank for hub-and-spoke airline	11
Figure 1.7	Example of a point-to-point network structure	12
Figure 1.8	Example of an aircraft route	14
Figure 1.9	Example of a crew trippair	15
Figure 2.1	Three hypothetical itineraries between Seattle (SEA) and Miami (MIA)	23
Figure 2.2	The characteristics of the hypothetical itineraries between Seattle (SEA) and Miami (MIA)	24
Figure 2.3	Sketch of a traveler's choice set	24
Figure 2.4	Three hypothetical itineraries between point A and point B	31
Figure 3.1	Example of network coverage for two competing airlines	40
Figure 3.2	Modeling framework for airline competition analysis and demand modeling	43
Figure 3.3	Example of a hypothetical network of three air carriers	44
Figure 4.1	Time-staged flight network	59
Figure 4.2	Time-staged flight network with aircraft assignment	60
Figure 4.3	Inbound and outbound flights at a hypothetical station for three aircraft types	61
Figure 4.4	Illustration of an interconnection node at a station	62
Figure 4.5	Illustration of the last interconnection node at a station	62
Figure 4.6	Illustration of the sizing constraints	64
Figure 4.7	Illustration of the aircraft count at the interconnection node	67
Figure 4.8	Example of a through flight	67
Figure 4.9	Example of three aircraft overnight at ORD and LAX	68
Figure 4.10	Example of a time-window arc	69
Figure 4.11	Example of interconnection nodes that connect the time-window arc	70
Figure 4.12	Example of fleet assignment with crew consideration	72
Figure 4.13	Example of a terminator flight	73
Figure 5.1	Example of an aircraft route	80
Figure 5.2	Example of a practical and efficient aircraft turn	80

Figure 5.3 Example of aircraft rotation 82
Figure 5.4 Example of through traffic 83
Figure 5.5 Representation of the aircraft routing solution 84
Figure 5.6 Representation of the aircraft routing solution with dummy variables 85
Figure 5.7 Many connection possibilities at the hub station compared to a spoke 86
Figure 6.1 Example of a typical crew trippair 90
Figure 6.2 Possibilities of crew connections at hub and spoke 92
Figure 6.3 Representation of the crew pairing solution 93
Figure 6.4 The trippairs matrix with the set of dummy trippairs 94
Figure 6.5 Example of the trippairs matrix, where each flight is only covered in one trippair 96
Figure 6.6 Example of the trippairs matrix, where each flight is only covered in one trippair, and the trippairs are sorted based on their cost 97
Figure 6.7 Selecting a subset T from the current trippairs matrix 97
Figure 6.8 Trippairs matrix with the hybrid approach 98
Figure 7.1 Example of trajectories of connecting passengers/baggage at airport terminal 110
Figure 7.2 Example of the flight assignment for a hypothetical set of gates 111
Figure 7.3 Example of eight flights assigned to gates 114
Figure 8.1 Sketch of the baggage-sorting facility 120
Figure 8.2 Illustrative example of the activity selection algorithm 123
Figure 8.3 General framework for the baggage-handling model 125
Figure 8.4 Example of trajectories of connecting baggage at the airport terminal 126
Figure 9.1 The framework of the flight planning process using the decomposition approach 133
Figure 9.2 Example of profile optimization 134
Figure 9.3 Examples of different fuel-loading patterns along the aircraft route 136
Figure 10.1 A sketch of the seats on a flight classified into full-fare seats and discounted seats 142
Figure 10.2 Booking pattern over time for business travelers and leisure travelers 143
Figure 10.3 Example of a demand-price curve for flight seats (one fare) 144
Figure 10.4 Example of a demand-price curve for flight seats (two fares) 144
Figure 10.5 Possible decisions on selling a seat on a flight 145
Figure 10.6 Seat availability for two booking classes 148

Figure 10.7	Representation of sequential nesting and seat availability for each booking class	149
Figure 10.8	Another graphical representation of sequential nesting	149
Figure 10.9	Example of a mix of parallel nesting and sequential nesting	150
Figure 10.10	Example of a point-to-point network structure	151
Figure 10.11	Example of a hub-and-spoke network structure	151
Figure 11.1	The flights of a hypothetical airline	156
Figure 11.2	Graphical representation of the different snapshots recorded for a hypothetical itinerary-fare class on a flight that is departing in 180 days' time	157
Figure 11.3	Graphical representation of the different snapshots recorded for a hypothetical itinerary-fare class on a flight that is departing in 90 days' time	157
Figure 11.4	Example of overestimated snapshots	158
Figure 11.5	Example of underestimated snapshots	158
Figure 11.6	Average origin-destination demand in the domestic US market	160
Figure 11.7	Example of a hypothetical itinerary	162
Figure 11.8	Representation of the last 52 observations of the demand	162
Figure 11.9	Most recent 52 observations of demand	163
Figure 11.10	A change in the schedule of the itinerary	164
Figure 11.11	Example of change of aircraft size	164
Figure 11.12	Probability distribution of demand	166
Figure 11.13	Example of two different probability distributions of demand with different levels of dispersion	166
Figure 11.14	Graphical representation of the expected seat revenue	170
Figure 12.1	Example of the probability distribution of the number of no-shows for a hypothetical flight	176
Figure 12.2	The relationship between the probability distribution of the number of no-shows and the number of overbooked tickets	178
Figure 13.1	Example of the probability distribution for the demand of the high-fare class	182
Figure 13.2	Graphical representation of the expected seat revenue	183
Figure 13.3	Expected seat revenue for each demand stream	185
Figure 14.1	Example of a hub-and-spoke network structure	190
Figure 14.2	A hypothetical air carrier network	195
Figure 15.1	The role of distribution channels for the air carriers	207
Figure 15.2	The current practice of air carriers' ticket distribution	209
Figure 16.1	Illustration of incremental revenue calculation	215
Figure 17.1	The routes of a hypothetical air carrier participating in multiple code-share agreements	222

Figure 18.1	Example of a ground delay program issued for a single flight	228
Figure 18.2	Example of a GDP issued for the flights arriving at a hypothetical airport	228
Figure 18.3	Another example of a GDP at a hypothetical airport where the adverse weather conditions extend over a longer period	229
Figure 18.4	Example of sharing information about canceling flight 5	231
Figure 18.5	Example of sharing information about delaying flight 5	232
Figure 18.6	Interaction between the FAA and the airlines during a GDP under the CDM paradigm	233
Figure 18.7	Example of substitution, where the air carrier that operates flights 4 and 5 exchanges their slots	234
Figure 18.8	Example of substitution, where the air carrier that operates flights 4 and 5 cancels flight 4 to move flight 5 into the vacated slot	235
Figure 18.9	Example of flight compression	236
Figure 19.1	A hypothetical aircraft route	242
Figure 19.2	Example of a trippair that has three duty periods and two layovers	243
Figure 19.3	Example of a crew misconnect problem	243
Figure 19.4	Example of crew rest (layover) problem	244
Figure 19.5	Example of crew duty problem	244
Figure 19.6	Example of a slack time of a resource	245
Figure 19.7	Resource slack with flight delay	246
Figure 19.8	Representation of the flight slack time	246
Figure 19.9	Connectivity under normal operation conditions	247
Figure 19.10	Resource connectivity after delaying flight F1	248
Figure 20.1	Resource connectivity after delaying flight F1	252
Figure 20.2	An example of two-way swap for two aircraft	253
Figure 20.3	An example of three-way swap for three aircraft	254
Figure 20.4	Aircraft is to return to its original schedule, if another swapping opportunity is found along its route	254
Figure 20.5	Hierarchy for the sequential recovery approach	257
Figure 20.6	The relationships between the main recovery desks in the airline's operations center	258
Figure 20.7	Flight delay as a recovery action	259
Figure 20.8	Forced flight cancellations due to GDP	262
Figure 20.9	Flight cancellation as a recovery action	262
Figure 20.10	Example of flight cancellation	263
Figure 20.11	Cancellation of round trip	263
Figure 20.12	Cancellation of a loop of three flights	264
Figure 20.13	Different cancellation plans in the aircraft route	264
Figure 20.14	Relationship between the flight delay, cancellations and resource recovery	265

List of Tables

Table 2.1	Examples of itinerary choice models	29
Table 2.2	The main characteristics of hypothetical itineraries	33
Table 3.1	The main characteristics of the different itineraries of the three air carriers	45
Table 4.1	Aircraft fleets for major US air carriers (Summer 2008)	54
Table 4.2	Flight range and seat capacity for most common aircraft fleets	58
Table 11.1	Probability distribution of the demand	167
Table 11.2	Probability of filling each seat	168
Table 11.3	Expected seat revenue	169
Table 12.1	The consequences of overestimating or underestimating the no-show rate	175
Table 13.1	Expected seat revenue	183
Table 14.1	The different itineraries served by flight DEN-ORD	190
Table 14.2	List of all origin-destination combinations served by the network	196
Table 14.3	Problem formulation using Excel solver	197
Table 14.4	Problem solution	199
Table 14.5	The dual values of the flight capacity constraints	200
Table 14.6	Calculations of itinerary contributions	201
Table 16.1	Expected demand and average fare for three hypothetical demand classes	216
Table 16.2	Calculation of expected seat revenue	217

This page has been left blank intentionally

Chapter 1
Introduction to Airline Management

Introduction

Aviation provides the only transportation network across the globe and it is crucial for global business development and tourism enrichment. Air transportation is one of the most important services to offer both significant social and economic benefits. By serving tourism and trade, it contributes to economic growth. It also provides jobs and increases tax revenues. Air transportation is essential for the fast movement of people and cargo shipments around the world. Finally, air transportation improves the quality of people's lives by broadening their leisure and cultural experiences. It gives a broad choice of holiday destinations around the world and is an affordable means to visit distant friends and relatives (ATAG 2005).

The use of commercial aviation has grown significantly over the last few decades, estimated to be more than seventy-fold since the first jet airliner flew in 1949 (ATAG 2005). This rapid growth is attributed to a number of factors. First, rising disposable income and quality of life in many parts of the world have encouraged more people in these areas to travel and explore opportunities overseas. Second, deregulation of aviation laws, and bilateral and open-sky agreements between governments have opened new markets for airlines, which make travel easier and cheaper. Third, demand is increasing because of growing confidence in aviation as a safe mode of travel. Fourth, increased efficiency and increasing competition have reduced world airfares and the cost of travel. Finally, globalization has increased the average distance traveled, as people do business in countries which now have improved political and social environments. The impact of these factors is expected to continue, however, at different levels in different parts of the world. The number of air travelers and the volume of air cargo is expected to continue to grow, increasing the pressure on all the contributors to the air transportation service to take advantage of opportunities and efficiently manage their service.

A major player in the air transportation industry is the airline. Current records indicate that there are more than 900 commercial airlines around the world, with a total fleet of nearly 22,000 aircraft (ICAO 2006). Commercial airlines serve nearly 1,670 airports through a route network of several million kilometers. These airlines transport close to 2 billion passengers annually and 40 percent of interregional exports of goods (by value). Also, an estimated 2.1 million people are employed by airline or handling agents: for example, as flight crew, check-in staff, and maintenance crew (ICAO 2006). Airline services are categorized as

being intercontinental, continental, regional, or domestic, and may be operated as scheduled services or charters. In terms of size, airlines vary from those with a single airplane carrying mail or cargo, through full-service international airlines operating many hundreds of airplanes. In many parts of the world, airlines are government-owned or supported. In recent decades, however, the trend has been to move toward independent, commercial public companies by giving more freedom to non-government ownership of airlines.

The increasing number of commercial airline companies has put more pressure on their management to continually seek profits, reduce cost, and increase revenues. Increasing demand for air transportation service has compelled airline management to take advantage of opportunities in different markets. At the same time, increasing competition among airlines necessitates that airline management seek efficiency in all their decisions to promote their profit. It is no surprise that many airlines throughout aviation history have been unable to remain in business, and in most cases, it is agreed that the demise of these airlines has been attributable to deficient management.

Airline management practice has evolved significantly over the past three decades. The development of this practice has contributed to recent advances in computation and communication technologies and, more importantly, the need to reduce costs and increase revenues. Nowadays airlines seek to perform efficiently in a competitive environment that only provides marginal profits. The airline business is characterized as being one of the most complex, involving multiple conflicting decisions that all need to be optimized at the same time. Several tactics have been developed and used to better plan and operate airlines. These tactics bank on scientific approaches available in operations research and mathematics literature to optimize airlines' decision-making processes, and are usually modeled within computerized systems that can automate decision making. Therefore, these scientifically-based tactics promise an easier decision-making practice for the airlines. The need for these tactics becomes more crucial as the size of the airline increases, and making decisions based on individuals' judgment or experience becomes more difficult. The next section highlights the main challenges of airline management that elaborate the complexity of the airline decision-making process.

Challenges of Airline Management

Impact of Other Players in the Industry

Airline management does not work independently of other players in the air transportation industry. Indeed, the decisions of airline management are very much affected by these other players. Figure 1.1 depicts the different entities that interact with airline management and affect decisions concerning government, airports, customers, alliances, suppliers, unions, and competitors.

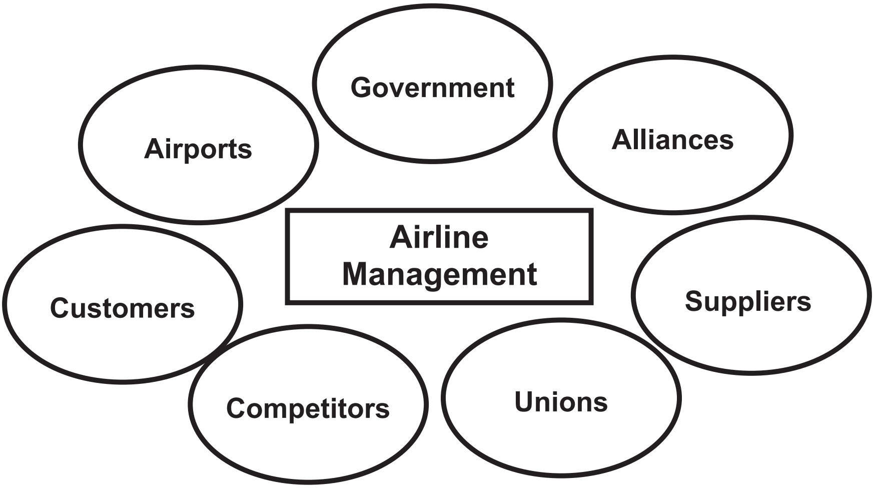


Figure 1.1 The different players in the air transportation industry

First, airline management must comply with the regulations of the airline's home country. It must also take into consideration and comply with the regulations of the governments of the countries where the airlines fly to and from, and whose airspace they cross. Governments typically watch competition between airlines and control airlines' strategic decisions, such as merging, acquisition between carriers, market entry or exit and pricing, environmental regulations, security regulations, maintenance, and safety requirements. Second, airline management should carefully consider the terms of their agreements with the different airports they serve. Several factors affect these agreements, including available infrastructure (gates, runways, baggage handling, and so on), expected traffic, airport charges and incentives, competition from nearby airports, available landing slots, congestion, and operational curfews. Third, an airline should consider the needs and preferences of its potential customers, the travelers. For instance, the airline should consider schedule convenience, competitive fares, onboard services, punctuality, and efficient customer service. Failing to fulfill the needs and preferences of customers might lead to losing them to other competing carriers or other modes of transportation. Fourth, in many cases, an airline participates in one or more alliances to expand its network coverage or share resources with other airlines. Several forms of alliance are available that reflect the level of cooperation between participating airlines. It is important for an airline to decide which alliance to participate in and how to share its resources efficiently with each member in the alliance to promote profitability. Typically, the airline has to maintain a certain level of operating standards to serve within a worldwide alliance.

Fifth, suppliers are crucial to the continuation of the airline's operation. Airlines depend on suppliers to provide important items such as aircraft, fuel, spare parts, meals, employee uniforms, and so on. Also, in many cases, airlines outsource to

vendors some of their jobs and services, such as aircraft maintenance, aircraft cleaning, ground handling, and sales. Therefore, an airline has to keep healthy relationships with its suppliers to continue operating successfully. Another entity in the air transportation industry that an airline has to deal with is unions. Different groups of workers form unions to achieve stronger negotiation power with airline management in terms of salary, benefits, or working rules. Keeping a good relation with labor in order to guarantee smooth operation of the business is one of the main objectives of airline management. Conflicts with unions might typically lead to negative actions by the unions, such as work slowdown or strikes, which usually impair the airline's operation significantly. Finally, in most markets, there is tough competition between several airlines. Typically, airlines continuously monitor the decisions of their competitors that relate to providing capacity, fare levels, fare restrictions, and departure times. In many situations, the decisions of the competing airlines proceed in a leader-follower pattern, where one airline takes an action and the other competing airlines try to find the best way to respond to this action.

Interacting Layers of Decisions

Like many other businesses, airlines management faces three levels of interacting decisions. These levels, as shown in Figure 1.2, include strategic, planning, and operations decisions. Strategic decisions typically require a long lead time before implementation and require a considerable monetary investment. They are also expected to have a significant impact on the form of the airline in the long term. Examples of strategic decisions include growth and expansion, fleet sizing (aircraft orders), hub locations, merging with other airlines, alliance participation, and location of maintenance facilities. Planning decisions are within a few months horizon, and can be defined as the process of efficiently using airline's available resources to maximize its revenue. The resources available to an airline include the facilities and the personnel that operate the business, including, for example, aircraft in different fleets, pilots with different qualifications, flight attendants, maintenance facilities, mechanics, gates, customer service agents, and ramp agents. The planning decisions include forecasting the demand between every origin-destination (OD), flight schedule development, assignment of flights to the different aircraft fleet (if the airline has more than one fleet type), aircraft routing across the different airports' with its maintenance consideration, planning the line of flight for pilots and cabin crew, crew accommodations, flight-gate assignment, and catering. Other planning decisions include the number of staff required to operate flights at different airports including customer service, ramp agents, baggage handlers, and so on. They also include decisions regarding fare levels in each OD market, fare restrictions, and seat inventory control for each flight. It should be mentioned here that these planning decisions are very dependent on each other, which makes the planning process complex.

The operations decisions for the airlines are those decisions that need to be verified or updated on an hourly or maximally on a daily basis. They include, for

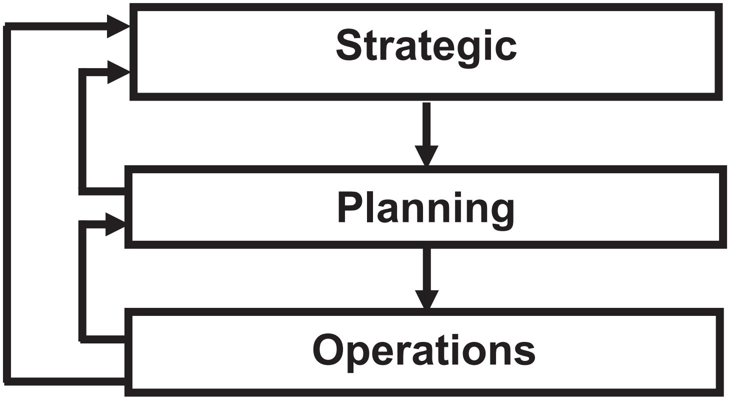


Figure 1.2 Decision levels of airline management

example, the response to unanticipated incidents such as adverse weather conditions, flights delays and cancellations, aircraft breakdown, and absence of crew or staff due to illness. Operations decisions also include watching revenues, bookings, and anticipated demand levels in the different markets, matching prices with competitors, and managing seat inventory on each flight on a daily basis.

Strategic decisions are expected to impact on planning decisions, which, in turn, affect the operations decisions. In addition, there is a reverse feedback from the operations phase to the planning phase, which also, in turn, may provide feedback to the strategic decisions phase. For example, the observation of a frequent delay of a certain flight waiting for its inbound aircraft might alert schedule planning to alter the schedule of this flight to give enough connection time for its inbound aircraft. Also, strong demand forecasting in markets might call for a change in the strategic plan regarding expansion and increase of fleet size. As explained in the next section, this book covers in detail the tactics currently practiced by airline management for the planning and operations phase. Strategic decisions are considered to be beyond the scope of this book.

Surrounding Events

The air transportation industry is characterized by the effects of rapid and significant impacts from surrounding events and economic and social changes. The negative impact on air transportation of factors such as wars, civil unrest, terrorist actions, increasing fuel prices, and epidemics has been clearly observed in several areas across the world. These events necessitate that airline management respond quickly and efficiently to study the impact of these events and take actions to alleviate their impact. To survive in business, in many situations, airlines may be forced to cut schedules, reduce fares, lay off employees, and cut salaries and benefits. For example, Figure 1.3 shows impacts on passenger demand, revenue, average fare, and average yield (revenue per seat mile) for airlines in the domestic

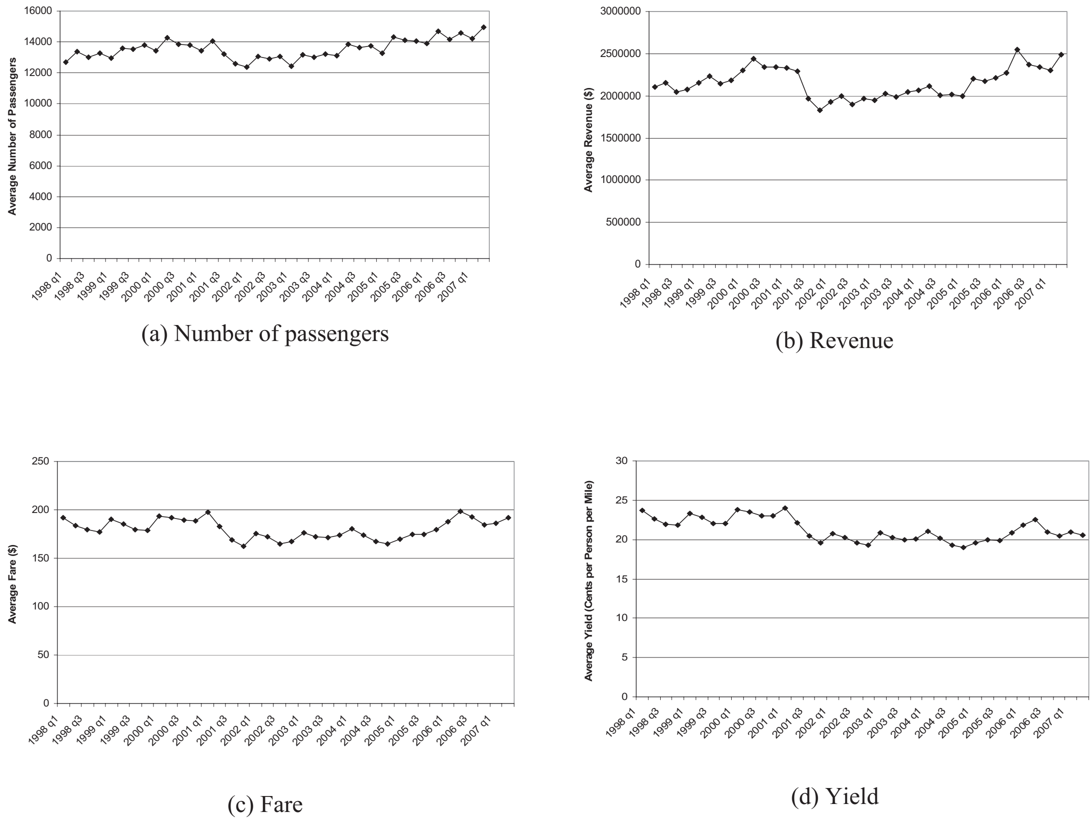


Figure 1.3 Changes in the main market characteristics due to the 9/11 terrorist attack in the domestic US market

US markets following the September 11, 2001 (9/11) terrorist attack. It is clear that these four measures were affected significantly because of this event. At that time, most domestic airlines considered significant actions such as cutting capacity, lowering fares, and discharging employees to respond to these market changes.

Many Groups to Contribute

Another challenge of airline operations is the interaction process among several groups of workers who work together to operate the flights. The product that an airline generates is a passenger seat or a space for cargo. This passenger seat or cargo space is typically a part of a flight that connects between two airports. The number of flights that an airline operates depends on the size of the airline. For large air carriers, the number of flights reaches a few hundred flights a day. Operating each flight requires significant cooperation among several groups of workers who all share the same objective of making the flight ready for departure on time. There are about 12 different groups who work on each flight before its departure. These groups include pilots (cockpit crew), flight attendants (cabin crew), maintenance crew, ramp agents, baggage-handling crew, cargo agents, fueling agents, customer service agents, gate agents, catering agents, aircraft cleaning agents, and operations agents or dispatchers. While the personnel in these groups differ in their qualifications, nature of work, workloads, and salary, they are all equally important for the departure of the flight. It is important for airline management to adequately set the work plan for each group, facilitate their work, and alleviate any possible conflict between them.

A pilot is a certified person who flies the aircraft of a certain aircraft fleet. Typically, each type of aircraft requires a certain number of pilots with certain specified qualifications. Flight attendants are airline staff employed primarily for the safety of passengers onboard. Their secondary function is the care and comfort of the passengers. The maintenance crew (maintenance) is responsible for servicing and repairing the aircraft to make sure that it is operational. Typically, maintenance performs several pre-specified mandatory service checks on the aircraft before departure, as specified by the manufacturers. Maintenance also performs several scheduled service checks on each aircraft in operation. Ramp agents help guide the aircraft to taxi in, park, and taxi out at the gate. Baggage handlers and cargo agents transport, load, and unload baggage and other cargo to and from the aircraft. Fueling agents provide fuel to the aircraft before departure or at intermediate stops in the flight. Customer service agents assist passengers with check-in, seat assignments, seat upgrades, and itinerary changes. Gate agents ensure that only authorized persons and passengers have access to the aircraft. Catering agents provide meals and beverages to be consumed on the flight. Aircraft cleaning agents clean the aircraft and the lavatories. The operations' agent or dispatcher coordinates the flight plan, weight, fuel requirements, and any weather-related or operations delays that are issued to the flight.

Airline Planning and Operations

As mentioned earlier, this book focuses on explaining the planning and operations phases of the airline. This section elaborates further on these two phases and introduces the different processes that are considered and explained in detail throughout this book. Figure 1.4 gives a sketch of the main processes considered in the planning and operations phases of the airlines. Planning starts by recording the anticipated demand and supply (available airline resources). Next, a set of interrelated planning processes is considered, including schedule planning, time banking, fleet assignment, aircraft routing, crew scheduling, airport facility planning, airport staff scheduling, pricing and seat inventory control, and sales and marketing initiatives. The planning processes are typically completed by a month to a few months before the implementation of the schedule, and they are repeated on a frequent basis as long as the airline is in business. The operations phase of the airline is concerned with implementing the planned airline schedule, while taking into consideration recovery for any unanticipated incidents such as adverse weather conditions, aircraft breakdown, crew absence, and so on. The operations phase is where decisions are made to recover the airline schedule from flight delay and cancellations, to compensate for missing or delayed aircraft and crew, and to reaccommodate stranded passengers. The operations phase also monitors seat bookings in the different markets and updates seat inventory control and pricing decisions. It should be mentioned that the current practice of airline planning and operations might differ to some extent, based on airline size and network structure. In the next subsection, the main objectives of each of the planning and operations processes are highlighted.

Network Structure

Airlines are typically classified as scheduled airlines or charter airlines. Scheduled airlines have a predefined flight schedule that is published through designated channels. In this schedule, the airline specifies the markets it flies to and the departure time and capacity of each flight in the schedule. Charter airlines, on the other hand, do not have a predefined schedule and typically operate on a demand basis. This book focuses on the business process of scheduled airlines, although many of these processes are also applicable to charter airlines.

Typically each scheduled airline has a predefined network structure. Selecting a network structure is considered one of the major strategic decisions of the airline. Most common network structures include 1) hub-and-spoke, 2) point-to-point, or 3) a combination of both. The hub-and-spoke network structure is one in which the airline considers one or more stations in the network to be its hub. Accordingly, any flight that is operated by this airline either starts or terminates at one of those hubs. The hub station is characterized by having numerous departures and arrivals every day for the airline. The spoke station has only a few departures and arrivals each

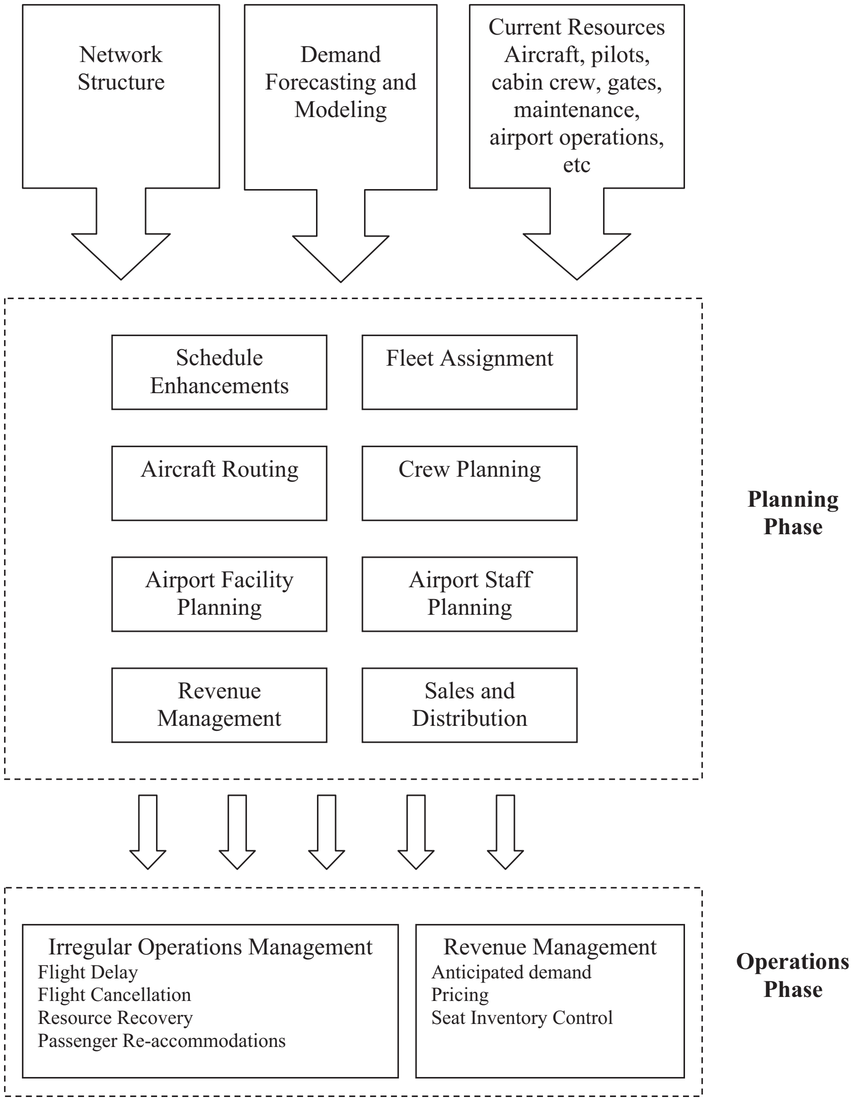


Figure 1.4 Processes considered in the planning and operations phases of the airlines

day. Figure 1.5 shows an example of a hub-and-spoke network for a hypothetical airline in the domestic US market that has a single hub at Dallas Forth-Worth airport, TX (DFW), and spokes at Atlanta, GA (ATL), Chicago, IL (ORD), Los Angeles, CA (LAX), San Francisco, CA (SFO), Denver, CO (DEN), Miami, FL (MIA), Newark, NJ (EWR), Detroit, MI (DTW), and New York, NY (JFK). This hub-and-spoke network structure gives considerable network coverage, enabling

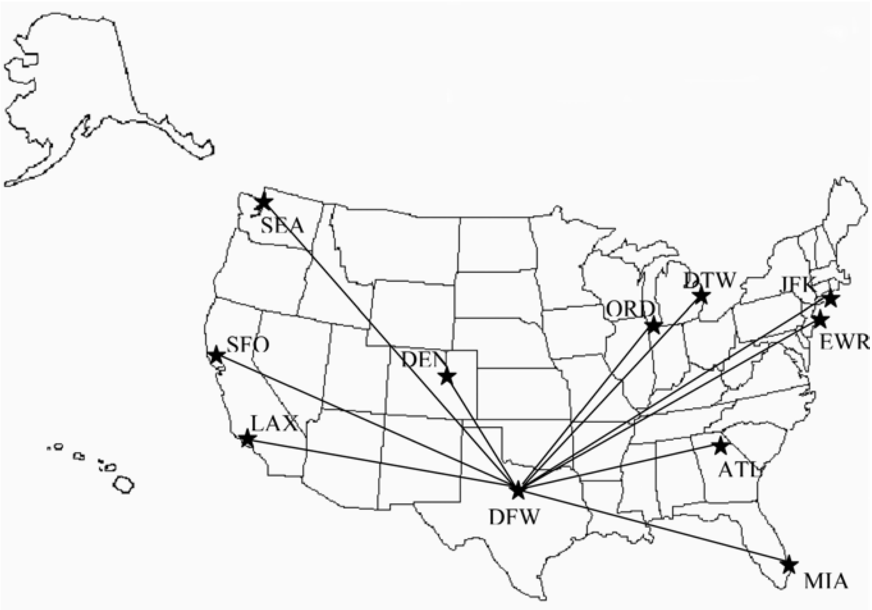


Figure 1.5 Example of a hub-and-spoke network structure

passengers to travel between any two cities served by the airline by physically connecting through the hub(s). Therefore, any flight in this network serves both local travelers between the origin and the destination of the flight and connecting passengers to and from other destinations.

For an airline adopting a hub-and-spoke network structure, to allow for smooth and convenient passengers connection at the hub, the arrivals and departures at the hub are adjusted in what is known as time banks. A time bank consists of a set of flight arrivals followed by a set of departures. Therefore, the time bank allows for several connection possibilities during a short period of time. For example, Figure 1.6 shows three hypothetical time banks at the DFW hub. In the first time bank, four flight arrivals (from SFO, SEA, LAX, and DEN) are connecting to four departures (to MIA, JFK, ATL, and EWR). Therefore, this time bank creates 16 different connection possibilities, as shown in Figure 1.6. Hub-and-spoke airlines pay considerable attention to selecting the flights included in each time bank to maximize passengers' connection possibilities and reduce unnecessary waiting time at the connecting hub. A hub-and-spoke network structure with condensed time banks, where many flights are scheduled to or from the hub over a short period of time, usually results in airport congestion. Airport congestion might lead to uncertainty in the taxi-in and taxi-out time of aircraft, which could affect the airlines' on-time performance. Several airports encourage the airlines to distribute their departures and arrivals over longer periods of time to alleviate this congestion; this is known as schedule de-peaking.

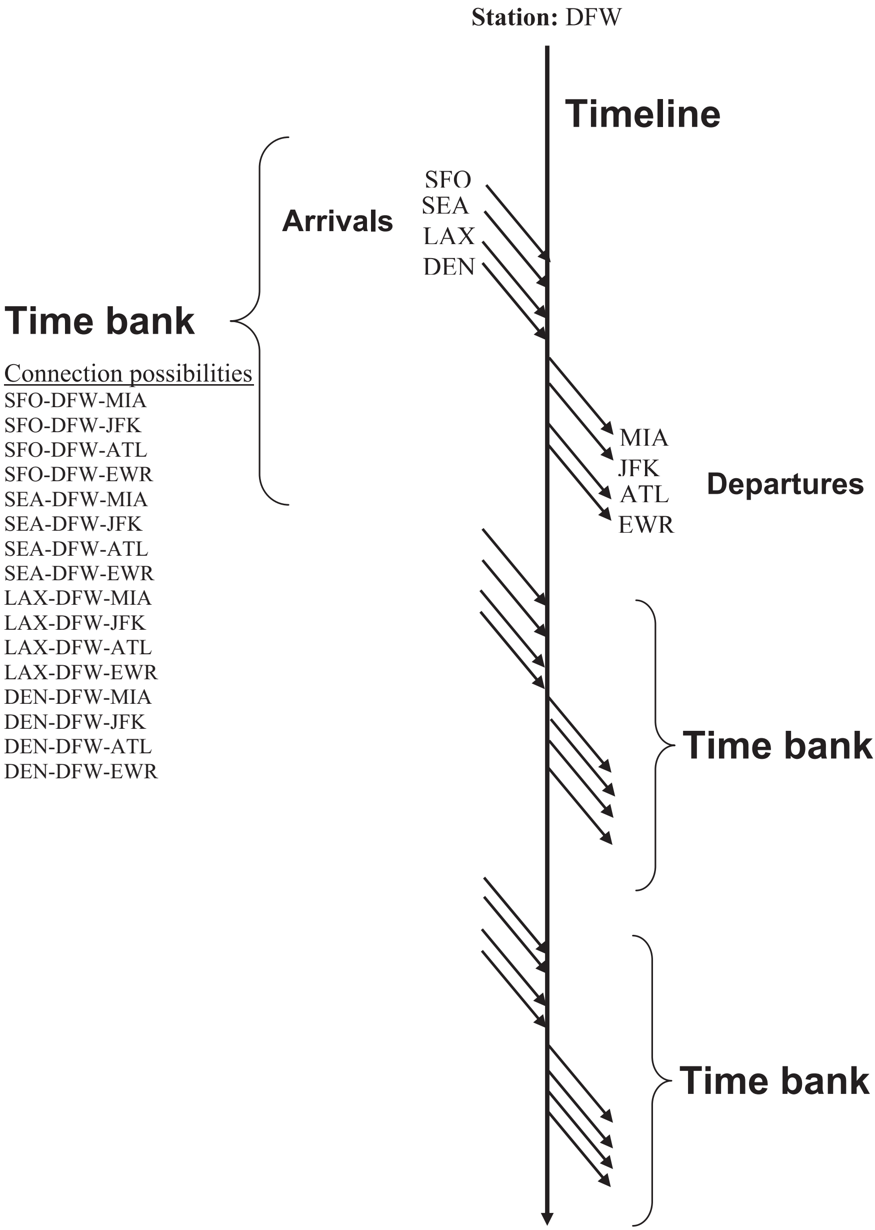


Figure 1.6 Example of time bank for hub-and-spoke airline

The hub-and-spoke network structure involves considerable interdependencies among the different resources of the airline, such as aircraft, pilots, and flight attendants. When an aircraft lands at the hub, it is typically scheduled to operate

another flight departing shortly from that hub. The same also applies to pilots and flight attendants. Accordingly, if a flight arrival is delayed at the hub, there is a chance that one or more departures will also be delayed while waiting for their inbound resources. Therefore, airlines that adopt a hub-and-spoke network structure must pay much attention to their flight on-time performance to minimize any downline impact of flight delay (also known as the snowball effect).

In the point-to-point network structure, as the name implies, airlines operate direct flights between cities. Therefore, the focus of these airlines is to serve local traffic between these two cities, and less attention is given to the connecting traffic beyond their immediate destinations. Because the point-to-point airlines do not depend on any connecting traffic to fill their flights, they have to select markets that have enough local demand, typically between large- and medium-size city pairs. The point-to-point airlines can also adjust arrivals and departures at one or more of their stations to allow for a few possibilities of profitably connecting itineraries for the passengers. Figure 1.7 shows a hypothetical example for a point-to-point network structure in the domestic US market.

Demand Forecasting and Modeling

Demand forecasting is the process of estimating the expected number of travelers on each flight in the schedule, given the flight schedules of all competitors in the

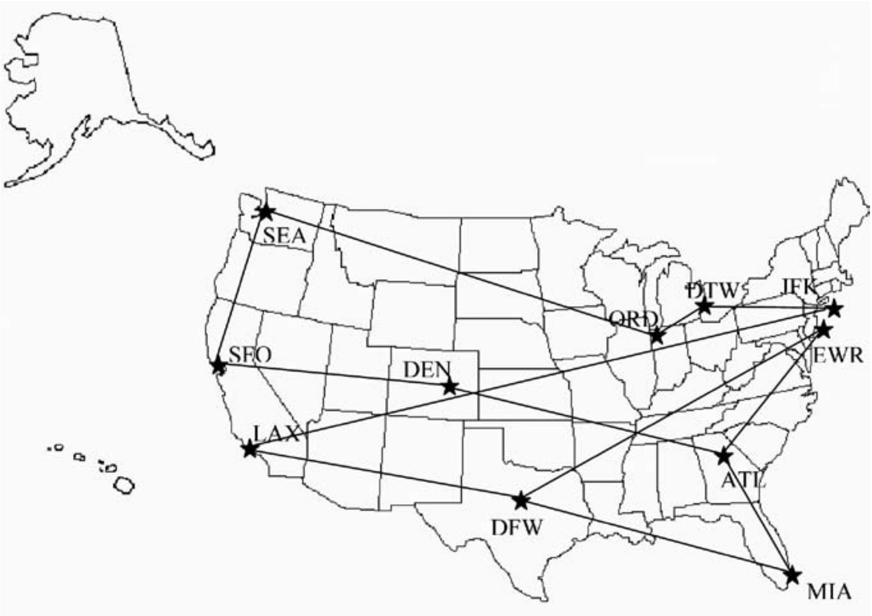


Figure 1.7 Example of a point-to-point network structure

different travel market. For example, Figure 1.5 shows the route map of a small airline operating in the domestic US market. This airline operates all its flights through a hub at DFW. As mentioned earlier, the different flights are scheduled such that passengers can travel between any two cities by connecting through the hub. For example, the passengers on the flight from Dallas (DFW) to Miami (MIA) are a mix of: 1) local travelers from Dallas to Miami, and 2) connecting travelers from other cities (San Francisco, Los Angeles, Seattle, and so on) to Miami. The demand forecasting and modeling process has to predict the passenger counts on each flight in the schedule and also estimate the possible changes in demand due to changes in schedule, pricing, competition, and so on. The demand forecasting process also estimates the airline share in each city-pair (OD market).

Fleet Assignment

The fleet assignment process is necessary for airlines that have more than one type of aircraft. It is the process of assigning the different flights in the schedule to the different fleet types. The process matches the characteristics of the aircraft and the flight to minimize the total cost of the flight to the airline. For example, the aircraft travel range must be consistent with the distance between the flight origin and destination. Also, in terms of the economics of fuel consumption, each aircraft type has an optimal range of travel distance in which it produces the best fuel consumption performance. Furthermore, the seat capacity of the aircraft should be consistent with the expected passenger count for the flight. Also, airport characteristics at the origin and the destination of the flight, including runway, gates, allowed noise levels, and curfews should be considered in the fleet assignment. Another logical constraint is maintaining continuity of fleet types at the different airport stations. The number of inbound flights assigned to a certain fleet type at any station should equal the number of outbound flights that are assigned to this fleet type. Other constraints related to the location of maintenance facilities and crew bases should be considered. Understandably there is a strong relationship between schedule design and fleet assignment. For example, an airline might decide to schedule and operate two small flights from point A to point B using two small aircraft or to schedule and operate one large flight using a large aircraft.

Aircraft Routing

Aircraft routing, as the name implies, means determining a route for each aircraft. Figure 1.8 presents an example of an aircraft route. As shown in the figure, the route of an aircraft consists of a sequence of flights and maintenance activities that extend over a few days (5–7 days). The flights are selected to ensure there is enough time between them to complete an aircraft turn or a maintenance activity. An aircraft turn is the time difference between the arrival time of a flight and the departure time of the next flight. An aircraft turn time should be long enough for

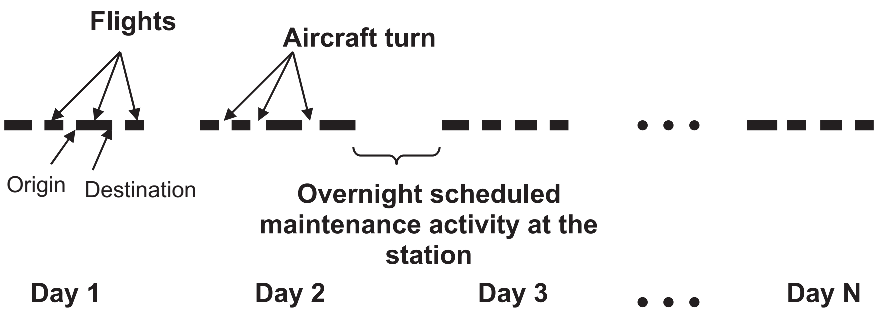


Figure 1.8 Example of an aircraft route

deplaning passengers of the inbound flight from the aircraft, unloading cargo and baggage, cleaning the aircraft, boarding passengers of the outbound flight, loading cargo and baggage, fueling, catering, exchanging crew, and so on. At the same time, the aircraft turn should not take long; long aircraft turns cause the aircraft to remain idle or unused over long periods of time. Longer idle time on the ground is expected to significantly dilute the revenue of the airline, as the throughput of its fleet declines.

Each aircraft must undergo certain maintenance activities, as specified by its manufacturer. Typically, the manufacturer specifies that a maintenance activity should be performed either after a few departures, certain flying hours, or certain operational hours. The maintenance activity of the aircraft extends over a few hours and is usually performed overnight at one of the airline's maintenance stations. Airlines usually position their maintenance facilities at one or more of their hub stations. Each aircraft must be scheduled for the required maintenance activities at the right time and at the right maintenance station.

Crew Planning

It is important for the crew (pilots and flight attendants) to know their traveling schedule ahead of time, so that they can plan their other life activities accordingly. The crew's work plan is typically extended over a period of one month. During the month, the crew member is classified either as a line crew or a reserve crew. A line crew gets a sequence of predefined trippairs over the month. A trippair represents a sequence of flights (segments), the first of which originates from the home city (domicile) of the crew and the last flight of which ends at the domicile. Figure 1.9 shows an example of a trippair for an airline crew. The trippair typically extends over two to seven days and consists of several duty periods followed by a rest (layover) period. A crew connection time occurs between every two successive flights in the same duty period, to enable the crew to connect from the gate of the inbound flight to the gate of the outbound flight. The duration of the minimum connection time, the maximum duty period, and the minimum layover period is pre-specified by the aviation authorities and the crew contract. Reserve crew

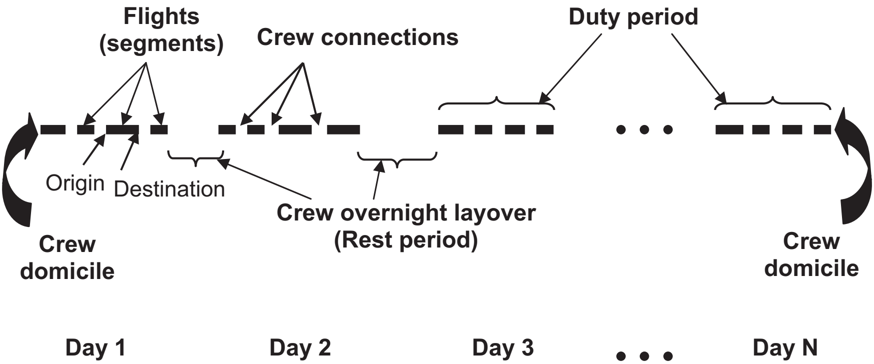


Figure 1.9 Example of a crew trippair

members are not given a line of flying. However, they are used as backup in case of system irregularity. It should be clear that the crew cost is one of the major cost components of the airlines. Efficient scheduling of the airline crew is expected to make considerable savings in airline costs.

Airport Facility and Staff Planning

During the planning phase, the scheduler must consider the different facilities at the different airports, including gates and baggage-handling facilities. The facilities at the airports should be planned efficiently to accommodate the planned flight schedule economically. The processes of facility planning are more crucial at the airline hub, where the airline operates many flights, and the usability of each resource is critical. The scheduling of airport facilities should take into consideration schedule disruptions that typically result from adverse weather conditions. Airport staff includes customer service agents, gate agents, baggage handlers, and ground agents. Typically, this staff works on shifts of about eight to ten hours every day. It is important for the airline to position the adequate number of staff at the appropriate times to operate the planned schedule economically and efficiently. The airport staff should also be adequate enough to manage any unplanned schedule disruptions at the airport due to adverse weather conditions or any other unforeseen factors.

Revenue Management

Airlines apply advanced revenue management (RM) techniques to maximize the revenue of their flights in different markets. RM is defined as selling the right seat to the right customer at the right price and at the right time. The idea behind RM is that travelers have different characteristics and primarily have different requirements for their travel. Basically, travelers can be classified as business travelers and leisure travelers. Business travelers are traveling for a work-related

trip or a business meeting. This group of travelers is typically less sensitive to the price of the ticket, because, in most cases, they are reimbursed for the cost of their travel by their employers. Business travelers have rigid travel plans that are typically constrained by predefined dates and times that usually span weekdays. They also tend to spend shorter periods of time at their destinations. They do not book their tickets far in advance and prefer flexible tickets that can be changed or canceled to match any possible changes in their travel plans. Leisure travelers, as the name implies, travel for recreational purposes or to visit family or friends. These travelers are sensitive to ticket prices. They also have flexible travel plans and tend to spend longer periods of time, including weekends, at their destinations.

Given that business travelers are more profitable to the airlines, the objective of RM is to ensure that enough seats are always available for these travelers, while the remaining seats on each flight are filled with low-revenue leisure passengers. The RM process involves three main modules, including pricing, demand forecasting, and seat inventory control. The main objective of pricing is to determine the right price for each market, taking into consideration competition from other carriers in the market. Demand forecasting means predicting the number of travelers by type in each market. Finally, the objective of seat inventory control is to assign seats on each flight to the different demand streams to maximize total revenue. The RM process is implemented for each future flight. The decisions of the RM process are updated on a daily basis until the day of the flight departure.

Sales and Distribution

Airlines expend considerable effort on sales and distribution initiatives that improve their market share and enhance profitability. These initiatives include, for example, relations with travel agents, global distribution systems, online ticket distribution channels, travelers' mileage plans, sales agreements with major businesses and promotions, and alliances and code sharing. Each of these initiatives needs proper evaluation in order to understand its impact on the airline profitability.

Irregular Operations Management

It is almost rare that an airline schedule is implemented as planned. Airline schedules are usually subject to disruptions due to adverse weather conditions, aircraft breakdowns, crew delays, and security breaches. When the airline schedule is disrupted, it is important for the airline to alleviate the impact of this disruption and recover the schedule in order to return to normal operations. When recovering the schedule, several objectives are considered by the airline. For example, the airline must minimize the deviation from the planned schedule by minimizing flight delays, cancellations, and crew swapping. In addition, it must not only adhere to the maintenance requirements of different aircraft at the right time, but also follow the regulations that govern the work rules of the crew on different

flights. Furthermore, the airline must comply with air traffic control regulations and programs that manage traffic in the airspace and at airports. Last but not least, it must minimize the total cost of recovery by avoiding expensive decisions such as flight cancellations, calling additional crew, and passenger rebooking on other airlines.

Structure of the Book

This book is structured in four main sections:

- Section I presents airline demand modeling and forecasting. In this section, we present the recent advances in modeling the airline network and the airlines' competition. Emphasis is given to itinerary choice models and factors affecting flight load factors and the airlines' market share.
- Section II is devoted to resource planning including aircraft, crew, and airport resources. This section presents recent advances in fleet assignment, aircraft maintenance routing, and crew scheduling. This section also discusses the recent advances in gate assignments and management of baggage-handling facilities. Techniques of flight planning and fuel management are also presented.
- Section III describes the process of airline RM, including demand forecasting and seat inventory control. It also offers an introduction to the ticket distribution practice, code sharing, and airline contract management with corporations to promote the number of business travelers.
- Section IV presents an introduction to the practice and tactics of irregular operations management. It provides an introduction to ground delay programs and the relation between the different actions in schedule recovery.

References

ATAG. 2005. *The Economic and Social Benefits of Air Transport*. Air Transport Action Group, 33 Route de l'Aéroport, P.O. Box 49, 1215 Geneva 15, Switzerland.

ICAO. 2006. *Economic Contribution of Civil Aviation*. International Civil Aviation Organization, 2006 Edition CD-ROM, Aviatech Publications.

This page has been left blank intentionally

SECTION I
Demand Modeling and Forecasting

This page has been left blank intentionally

Chapter 2
Modeling the Choice of Travel Options

Introduction

On a given day, a city pair (for example, Miami-Seattle, Tokyo-Frankfurt) is served by hundreds of flight itineraries. Each of these itineraries has different characteristics related to price, the number of connections, connection durations, departure times, arrival times, and the total travel distance. Recently, airlines have paid more attention to understanding how travelers select among their air travel options (Coldren et al. 2003), investigating in particular, what factors affect travelers' decisions on itinerary choice and the extent of the effect of each factor. Understanding the impact of these factors is crucial to the airlines in planning their itineraries in a way that is attractive to travelers. This understanding also allows airlines to estimate the market share of their itineraries in any city pair under different competition scenarios.

A reasonable approach to investigating the impact of itinerary characteristics on itinerary attractiveness is to study historical data. Historical data reveal the actual behavior of travelers regarding how they select their travel itinerary among different available itineraries that have different characteristics. Historical data can be used to develop mathematical models that represent how travelers select their travel itinerary. A mathematical model is defined as a mathematical representation of reality that attempts to explain the behavior of some aspect of the real situation. It is a process that relates a number of variables that are defined to represent the inputs and outputs of the process. The mathematical model serves the following purposes: 1) to extrapolate historical data to derive meaning; 2) to establish an understanding of the relationships among the input data items within a model; and 3) to answer what-if questions related to the modeled system. To develop a mathematical model that represents any system there should be a reasonable understanding of how this system works. In addition, the factors that affect this system should be identified. In the next few sections, an introduction is given to the itinerary choice process and how this process can be modeled mathematically. An example from the industry for the itinerary choice models is also presented.

Factors Affecting Itinerary Choice

Factors that affect the traveler's itinerary choice can be classified into three main categories. The first category includes the factors related to the characteristics of the available itineraries, which include price, departure and arrival times, number

of connections, connection duration, overall travel distance, aircraft type and size, and air carrier reputation. It is generally expected that travelers prefer to select cheap itineraries, fewer connections, convenient connection periods and locations, convenient departure and arrival times, less travel distance, and reputable airlines. Therefore, itineraries that enjoy these characteristics are likely to be selected by travelers.

The second category is related to the individual's socioeconomic characteristics, which include income, age, gender, and participation in the airline's loyalty program(s). An individual's income affects sensitivity to ticket prices. Travelers with a high income are usually less sensitive to price cost. A traveler's age and gender also affect the choice of travel option. For instance, older ladies might want to avoid arriving late at night at a destination, so they may avoid itineraries with late arrival times. Also, they may perceive itineraries with connections very negatively compared to nonstop itineraries. Travelers usually participate in airline loyalty programs (frequent flyer program or mileage plan program), where they receive points for every mile of travel with the airline. These miles can be redeemed for free flights or other travel amenities from the airline. These loyalty programs make travelers more willing to select itineraries that belong to the airline(s) with the loyalty program.

The third category is pertinent to trip characteristics, which could be for business or leisure, domestic or international. This category also includes the geographical location of the origin and destination of the trip. The trip purpose, such as business or leisure, affects the itinerary choice. For example, business travelers are usually less sensitive to ticket price than leisure travelers. Conversely, they are more sensitive to itinerary schedule. Business travelers usually prefer to travel after the end of the business day with nonstop itineraries. Generally, travelers have different perceptions of itineraries based on the length of haul and whether the trip is domestic or international. For example, travelers may be more willing to accept connecting itineraries with long layover (connection) periods when they are traveling long-haul international trips. For short-haul domestic trips, connecting itineraries with long layover periods are usually not acceptable. Also, the geographical location of the origin and destination, including time zone differences, might affect itinerary selection. The acceptance of the red-eye overnight flights from the west coast to east coast of the US, which arrive early in the morning at the east coast destination, is an example of the impact of time zone on travel preferences.

Itinerary Choice Problem Formulation

It is assumed that an individual i is traveling from origin station O to destination station D. There is a set of itineraries J that connect the city pair (OD). Each itinerary $j \in J$ has different characteristics including price, connection quality,

departure and arrival times. Individual i has to evaluate the available itineraries in set J and select the most attractive itinerary.

To further illustrate the situation, consider the hypothetical case in the domestic US market of an individual traveling from Seattle (SEA) to Miami (MIA). Assume also that there are three itineraries that connect the city pair SEA-MIA, as shown in Figure 2.1. The first itinerary connects through Chicago O'Hare airport (ORD). The second itinerary connects through Dallas-Fort Worth airport (DFW). The third itinerary is a nonstop itinerary. Figure 2.2 shows the main characteristics of the three itineraries, including the departure and arrival times of each flight and the price. Therefore, an individual traveling from SEA to MIA selects only one alternative from a choice set that includes three different alternatives as shown in Figure 2.3.

Next, we should consider how individual i selects an itinerary. From a close look at the characteristics of the three itineraries, it can be concluded that the first itinerary is the cheapest. However, it has a very early morning departure and one connection at ORD. Conversely, the third itinerary is the most expensive. However, it is a nonstop itinerary with a convenient departure time. Traveler i evaluates all the itineraries and trades off their characteristics. For instance, traveler i might trade schedule convenience for price. If schedule convenience is worth the difference in the price between the first and third itinerary (\$620–\$490), the traveler might select the third itinerary. On the contrary, traveler i may select the cheapest itinerary, if the schedule convenience is not worth the price difference.

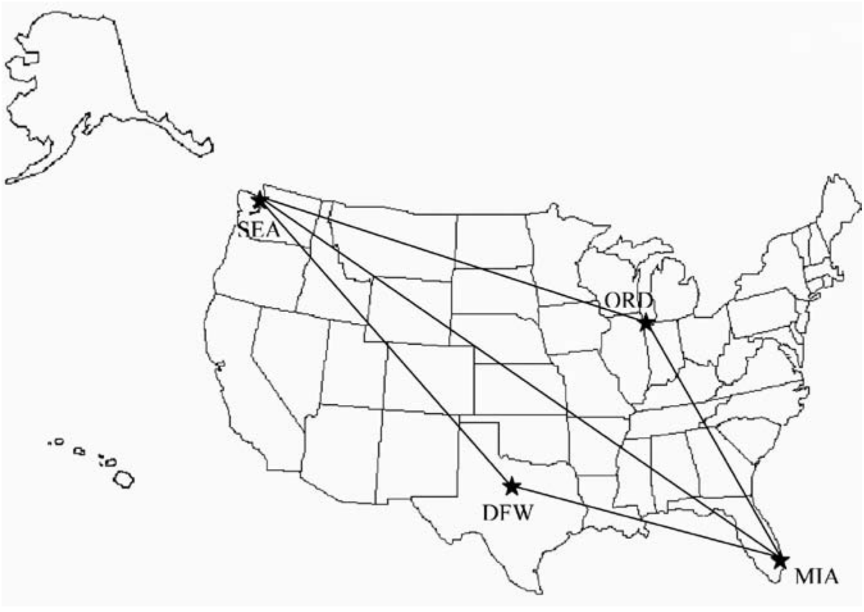


Figure 2.1 Three hypothetical itineraries between Seattle (SEA) and Miami (MIA)

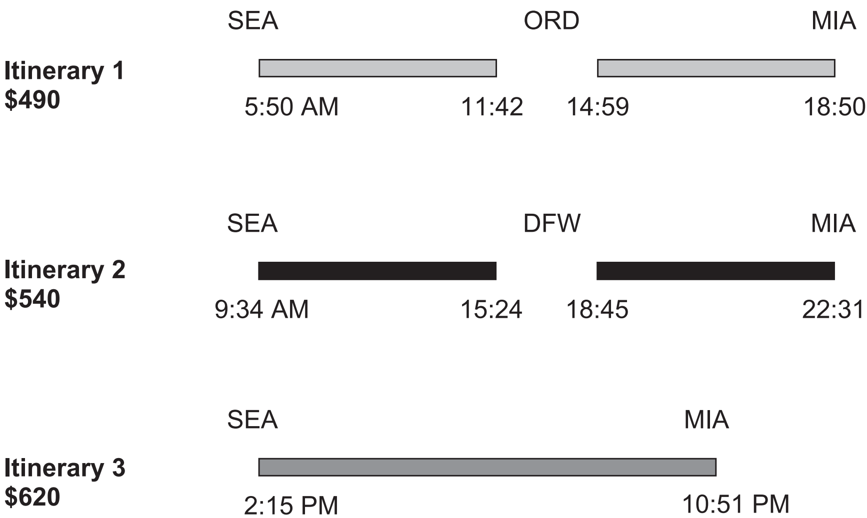


Figure 2.2 The characteristics of the hypothetical itineraries between Seattle (SEA) and Miami (MIA)

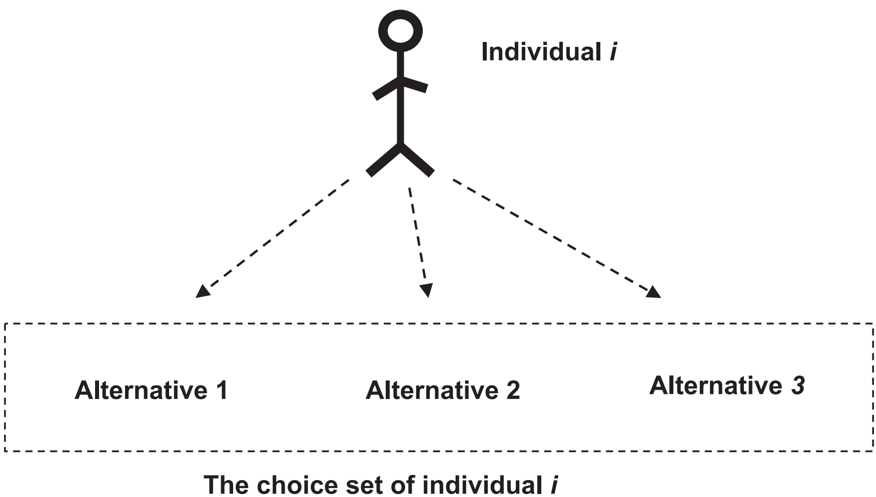


Figure 2.3 Sketch of a traveler's choice set

The next logical question is to quantify how individual travelers trade off the different characteristics of the different alternatives. That is, what itinerary characteristics are valuable to travelers when they are selecting their travel options. The discrete choice theory, which is presented briefly in the next section, provides an answer to this question.

Discrete Choice Theory

A quantitative term U_{ij} is defined to measure the attractiveness (utility) of individual i with respect to itinerary j. It can be represented as a score that individual i is given to alternative j. The higher the value of the utility for an itinerary, the more attractive this itinerary is for the traveler. Also, as the value of the utility of an itinerary increases, the chance of selecting this itinerary by traveler i is expected to increase. The utility, U_{ij} can be represented as a function of: 1) characteristics of itinerary j; 2) characteristics of traveler i; and 3) trip characteristics. The utility function can be represented mathematically as follows:

$$U_{ij} = f\text{(itinerary characteristics, traveler characteristics, trip characteristics)} \quad (2.1)$$

The utility function is composed of two main components; the deterministic (measurable) component and the random component.

$$U_{ij} = V_{ij} + \varepsilon_i \quad (2.2)$$

Where:
V_{ij} = the measurable component of the utility function
ε_i = the random component of the utility function.

The deterministic component can typically be explained and measured by the characteristics of the itinerary, the traveler, and the trip. The component generally measures how the utility of an alternative changes with the changes of the characteristics of this alternative. The random component represents the unexplained behavior of travelers that might be attributed to errors, change in taste, and other unknown factors that contribute to the evaluation of the alternatives.

To further explain these two components, consider the case where a traveler is selecting between two identical itineraries, but one is cheaper than the other. For a rational traveler, the cheaper itinerary will have a higher utility (attractiveness) compared to the expensive one. Also, there is high chance that the cheaper itinerary will be selected by the traveler. In this example, the difference in utility can be attributed to, and explained by, the difference in price between the two itineraries. Given this example, if a traveler selects the more expensive itinerary, further explanation is needed. For this traveler, the utility of the expensive itinerary is higher than the utility of the cheaper one. This can be attributed to several factors. For example, this traveler might have made an error when evaluating the alternatives. Also, there could be other unknown factors that affected the attractiveness of the alternatives other than the price. A random component is considered in the utility function of the model to take care of these situations.

The measurable (deterministic) component of the utility function is typically estimated using historical data. Because the random component cannot be estimated, an assumption should be made regarding how the random component is statistically distributed. The random component can be distributed according to any probabilistic distribution such as a normal distribution or Gumbell distribution. In the case that the error component is distributed according to the Gumbell distribution, the probability of selecting an alternative *j* by an individual *i* can be estimated using the logit model, which can be represented mathematically as follows:

$$p_{ij} = \frac{e^{V_{ij}}}{\sum_{j \in J} e^{V_{ij}}} \tag{2.3}$$

Where:
p_{ij} = the probability that individual *i* is selecting alternative *j*
V_{ij} = the measurable component of the utility function.

A more comprehensive review of multinomial logit models can be found in McFadden (1978), Ben-Akiva and Lerman (1985), and Koppelman and Sethi (2000).

In the next two sections, an industry example for estimated utility function and an illustrative example for the itinerary choice models are presented respectively.

Example of Itinerary Choice Models (Coldren et al. 2003)

The work by Coldren et al. (2003) represents one of the rigorous attempts to model itinerary choice in the domestic US airlines' market. In this work, historical ticket booking data in the US domestic market are collected and used for this purpose. These data represent the individual bookings by most major US airlines. Each record in these data represents an itinerary that is selected by a traveler. Each itinerary is defined by its main characteristics, including marketing and carrying airlines, departure and arrival time, number of stops, connection quality, aircraft type, and price. Individual characteristics such as gender, age, and income are not considered because this information is not available in the booking data. The booking data are categorized based on the geographical locations of the origin and destination of the trip. Four main regions are considered for the trip origin and destination according to the time-zone classification in the domestic US market. These regions include East (E), Central (C), Mountain (M), and West (W). Therefore, the booking data are categorized into 16 main categories based on the origin and destination of the trips (E-E, E-W, E-M, E-C, W-W, W-E, and so on). For example, the E-W category includes itineraries with the origin being one of

the cities in the east coast region (eastern time zone), and the destination being one of the cities in the west coast region (Pacific time zone). In addition to these 16 categories, two other categories are considered: trips from the US mainland to Alaska and Hawaii, and from Alaska and Hawaii to the US mainland. Therefore, in total, booking data are categorized into 18 classes based on the geographical locations of the origin and destination of the trip.

Next, ticket booking data in each geographical category are used to estimate a special logit model to represent itinerary choice in the markets that belong to this category (the estimation process is beyond the scope of this book). Accordingly, 18 different logit models are developed from these data. (Variables that explain the itinerary choice behavior of travelers that are used in the models are given below.) Table 2.1 gives the different explanatory variables used in the logit models. It also gives the values of the parameters for the East-East (E-E) and West-East (E-W) cases. Other models for other regions can be found in Coldren et al. (2003).

As shown in Table 2.1, the variables that are included to explain itinerary choice are all related to itinerary characteristics. Travelers' characteristics (for example, income, age, and gender) are not included because they are not available in the bookings dataset. The considered explanatory variables are classified into level-of-service, connection quality, carriers' attributes, aircraft size and type, and time of day.

The level of service variable is a dummy (0-1 binary) variable that represents the connection quality of the itinerary with respect to the best itinerary available in the city pair. An itinerary can be nonstop, direct, single-connect, or double-connect. A nonstop itinerary, as the name implies, consists of one flight only. A direct itinerary is when the aircraft lands at an intermediate airport to drop off and pick up additional passengers to the final destination. Passengers of the direct itinerary do not switch aircraft and gate at the intermediate airport. A single-connect itinerary is composed of two flights that connect at two different gates at the connecting station. Finally, a double-connect itinerary is composed of three flights. Itineraries with more than two connections are not considered, because these itineraries are not common for travel in the domestic US market. The different markets are classified as nonstop, direct, or single-connect. A market is defined according to the best itinerary serving in this market. A nonstop market is a market that has at least one nonstop itinerary. For instance, a city pair such as Atlanta-Dallas-Fort Worth (ATL-DFW) is a nonstop market because there is at least one nonstop itinerary that connects between ATL and DFW. A single-connect market is a market that has at least one single-connect itinerary and at the same time is not served by any nonstop or direct itineraries. The best itinerary available in this market is the single-connect itinerary. For instance, a city pair such as Daytona Beach-Seattle (DAB-SEA) is a single-connect market because the best itinerary that connects between these two cities is a single-connect itinerary.

Following the discussion given by Coldren et al. (2003), the parameters for the level of service variables indicate that passengers prefer itineraries with fewer

connections. As the number of connections in the itinerary increases, the total utility (attractiveness) of the itinerary decreases. For example, for the E-E model, if the market is a nonstop market, the double-connect itinerary has a parameter of -7.7557 compared to -3.1213 for the single-connect itinerary, -1.6819 for the direct itinerary, and zero for the nonstop itinerary (which is used as a reference point). In addition, the estimated parameters indicate that the same itinerary type is perceived differently based on the market type. For instance, the double-connect itinerary has a parameter of -7.7557 in the nonstop market, while the parameter of the double-connect itinerary has a parameter of -3.0953 in a single-connect market. This means that double-connect itinerary is relatively more acceptable in single-connect markets than in nonstop markets.

The second group of variables is related to the connection quality. When travelers are connecting at any airport, they prefer to connect through the best connection in the airline's schedule. The second-best connection itinerary is perceived as less than the best connection itinerary, as indicated by the corresponding parameter in Table 2.1. Also, as the time difference between the best connection and the second best connection (second-best connection time difference) increases, the second-best connection itinerary becomes less attractive. For example, for the E-E model, the second-best connection variable has a negative parameter of -0.556. Also, for each minute of the second-best connection time difference, the utility of the itinerary changes by -0.0157.

Next, the best connection time difference is the elapsed time between an itinerary involving a connection and the fastest itinerary involving a connection for each city pair independent of transfer cities. The longer the connection period of an itinerary compared to the fastest connecting itinerary, the less attractive the itinerary becomes, as indicated by the negative sign of the corresponding parameter in the two models presented in Table 2.1. Similarly, the models indicate that travelers prefer itineraries with the least travel distance, which is logical.

The amount of presence of an airline in the origin or the destination of a trip impacts on the attractiveness of the itineraries offered by this airline. The model shows that travelers prefer the itineraries of airlines that have a heavy presence at the origin or the destination of the trip. This preference is indicated by the positive value of the parameter corresponding to the point-of-sale-weighted city presence. For instance, travelers to and from Atlanta (ATL) are more likely to buy tickets for Delta Airlines because Delta Airlines has a high presence in Atlanta. Similarly, travelers who travel to and from Dallas-Fort Worth (DFW) are more likely to buy tickets for American Airlines because American Airlines has a high presence at DFW.

Logically, and as shown by the model, expensive itineraries are less attractive to travelers. The model also shows that some air carriers are more preferable than others. However, this information is not presented in the work of Coldren et al. (2003). The model also shows that code-share itineraries are less attractive than itineraries that are marketed and operated by the same airline. For example, in the

Table 2.1 Examples of itinerary choice models (Coldren et al. 2003)

Explanatory Variables	E-E	E-W
Level-of-Service		
Nonstop Itinerary in Nonstop Market	0	0
Direct Itinerary in Nonstop Market	-1.6819	-1.6615
Single-Connect Itinerary in Nonstop Market	-3.1213	-2.9729
Double-Connect Itinerary in Nonstop Market	-7.7557	-7.213
Direct Itinerary in Direct Market	0	0
Single-Connect Itinerary in Direct Market	-0.7901	-1.0615
Double-Connect Itinerary in Direct Market	-4.9442	-4.4475
Single-Connect Itinerary in Single-Connect Market	0	0
Double-Connect Itinerary in Single-Connect Market	-3.0953	-2.8209
Connection Quality		
Second-Best Connection	-0.556	-0.7269
Second-Best Connection Time Difference	-0.0157	-0.0162
Best Connection Time Difference	-0.0108	-0.0104
Distance Ratio	-0.0125	-0.0173
Carrier Attributes		
Point of Sale Weighted City Presence	0.0024	0.0071
Fare Ratio	-0.0018	-0.0035
Carrier Constants (Proprietary)	xxxxxx	xxxxxx
Code Share	-1.5911	-2.1658
Aircraft Size and Type		
Propeller	-1.242	-0.9988
Regional Jet	-0.7046	-0.7079
Propeller Seats	0.0246	0.0184
Regional Jet Seats	0.0117	0.008
Mainline Seats	0.0041	0.0032

Table 2.1 ***Concluded***

Explanatory Variables	E-E	E-W
Time of Day		
Midnight–5 AM	-1.1653	-1.2037
5–6 AM	-0.4653	-0.5037
6–7 AM	0	0
7–8 AM	0.2865	0.2297
8–9 AM	0.2836	0.3168
9–10 AM	0.2046	0.3413
10–11 AM	0.1219	0.3425
11–12 noon	0.1022	0.267
12–1 PM	0.1452	0.2977
1–2 PM	0.1732	0.2594
2–3 PM	0.2622	0.2462
3–4 PM	0.3438	0.2557
4–5 PM	0.423	0.2097
5–6 PM	0.4667	0.2324
6–7 PM	0.4054	0.1889
7–8 PM	0.2047	0.0472
8–9 PM	-0.0362	-0.2243
9–10 PM	-0.3565	-0.3924
10–Midnight	-0.6468	-0.317

E-E model, a negative sign of -1.5911 is obtained for the parameter of the code-share dummy variable.

The estimation results of the model show that travelers prefer large (mainline) jets to regional jets and region jets to propeller aircraft. Also, larger aircraft are preferred over smaller aircraft (within an aircraft type). This preference is mainly due to the perceived levels of comfort and safety. The parameter estimates for aircraft type and number of seats have the correct sign and are of reasonable magnitude.

The last group of variables is related to the departure time of the itinerary. The departure time is considered as a dummy variable for each hour of the day (based on the departure time of the first leg of the itinerary). As shown in the presented models, the parameters of some departure time variables have a negative sign, which means

that those departure times are less attractive to travelers. For example in the E-E model, early morning (before 6:00 AM) and late night (after 8:00 PM) departure times are less desirable compared to other departure times. For the E-E model, the most attractive departure time during the morning period is 9:00–10:00 AM. The corresponding parameter for this variable has the highest value (0.3204) among the parameters of other departure times. Similarly, the most attractive departure time during the afternoon period is 5:00–6:00 PM. This departure time may be suitable for business travelers, who prefer to travel after the end of the business day.

Illustrative Example

In this section, a simplified illustrative example is presented to show how the itinerary choice models can be used to represent the itinerary choice behavior of travelers and estimate the market share of different competing airlines. Consider a hypothetical domestic US market, where the origin (point A) is in the east coast region and the destination (point B) is in the west coast region. There are two different airlines competing in this market. The first airline offers one nonstop itinerary, while the second airline provides two single-connect itineraries, as shown in Figure 2.4. The characteristics of the three itineraries are given in Table 2.2. For example, as shown in Table 2.2, itinerary 1 is a nonstop itinerary, which causes the market AB to be defined as a nonstop market. The departure time of

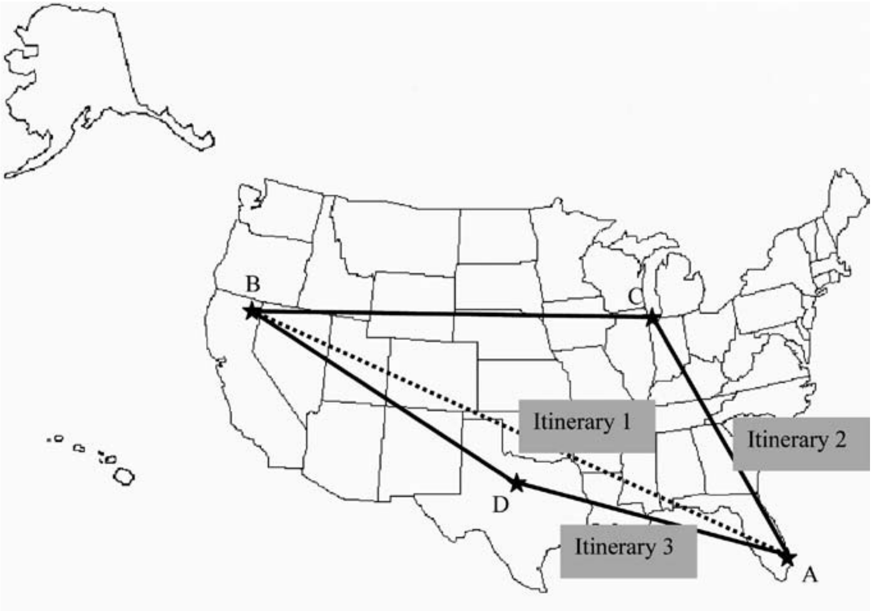


Figure 2.4 Three hypothetical itineraries between point A and point B

itinerary 1 is at 5:00–6:00 AM. Itinerary 1 has the shortest distance between A and B, with distance ratio of 100 and its cost is twice the market average (fare ratio is 200). Airline I that operates itinerary 1 has a 30 percent city presence and an aircraft capacity of 200 seats.

First, the utility function of each itinerary is calculated. The utility is calculated by the summation of the multiplication of each variable considered in the model by its corresponding parameter. The corresponding parameters are obtained from the E-W model of Table 2.1. The utilities of the three itineraries are calculated as follows:

The utility of itinerary 1:

$$U_1 = (-0.0173)(100) + (0.0071)(30) + (-0.0035)(200) + (0.0032)(200) - 1.2037 = -2.0807$$

The utility of itinerary 2:

$$U_2 = -2.9729 + (-0.0173)(108) + (0.0071)(30) + (-0.0035)(70) + (0.0032)(150) + 0.0311 = -4.3622$$

The utility of itinerary 3:

$$U_3 = -2.9729 - 0.7269 + (-0.0612)(15) + (-0.0104)(100) + (-0.0713)(105) + (0.0071)(40) + (-0.0035)(100) + (0.0032)(150) + 0.2324 = -5.2169$$

Second, using equation 2.3, the probability of selecting each alternative is calculated as follows:

The probability of selecting itinerary 1:

$$P_1 = \frac{e^{-2.0807}}{e^{-2.0807} + e^{-4.3622} + e^{-5.2169}} = 0.873 = 87.3\%$$

The probability of selecting itinerary 2:

$$P_2 = \frac{e^{-4.3622}}{e^{-2.0807} + e^{-4.3622} + e^{-5.2169}} = 0.089 = 8.9\%$$

The probability of selecting itinerary 3:

$$P_3 = \frac{e^{-5.2169}}{e^{-2.0807} + e^{-4.3622} + e^{-5.2169}} = 0.038 = 3.8\%$$

According to the estimated probability, there is high probability, 87.3 percent, that an average traveler will select itinerary 1 for their travel from

Table 2.2 The main characteristics of hypothetical itineraries

Variables	Itinerary 1	Itinerary 2	Itinerary 3
Level-of-Service			
Nonstop Itinerary in Nonstop Market	1		
Direct Itinerary in Nonstop Market			
Single-Connect Itinerary in Nonstop Market		1	1
Double-Connect Itinerary in Nonstop Market			
Direct Itinerary in Direct Market			
Single-Connect Itinerary in Direct Market			
Double-Connect Itinerary in Direct Market			
Single-Connect Itinerary in Single-Connect Market			
Double-Connect Itinerary in Single-Connect Market			
Connection Quality			
Second-Best Connection			1
Second-Best Connection Time Difference			15 (minutes)
Best Connection Time Difference		0 (minutes)	10 (minutes)
Distance Ratio	100	108	105
Carrier Attributes			
Point of Sale Weighted City Presence	30	30	40
Fare Ratio	200	70	100
Carrier Constants (Proprietary)			
Code Share			
Aircraft Size and Type			
Propeller			
Regional Jet			
Propeller Seats			
Regional Jet Seats			
Mainline Seats	200	150	150

Table 2.2 *Concluded*

Variables	Itinerary 1	Itinerary 2	Itinerary 3
Time of Day			
Midnight–5 AM			
5–6 AM	1		
6–7 AM			
7–8 AM			
8–9 AM			
9–10 AM		1	
10–11 AM			
11–12 noon			
12–1 PM			
1–2 PM			
2–3 PM			
3–4 PM			
4–5 PM			
5–6 PM			1
6–7 PM			
7–8 PM			
8–9 PM			
9–10 PM			
10–Midnight			

point A to point B. The probabilities of selecting itineraries 2 and 3 are 8.9 percent and 3.8 percent, respectively. The sum of the three probabilities should equal 100 percent. In other words, the probability of selecting airline 1 is 87.3 percent, and the probability of selecting airline 2 is 12.7 percent (8.9 percent plus 3.8 percent). These calculated probabilities are sometimes referred to as the unconstrained market share of the airline, which is the demand share of the airline when there is unlimited seating capacity on its itineraries. In the next chapter, we will explain how the itinerary choice models can be used within a demand modeling framework to estimate the number of passengers on different flights and the airline's share in different markets.

Primary Contributions

There has been other research that has investigated modeling air travelers' choice behavior. For example, Nason (1981) develops a multinomial logit model for airline choice by using the stated preference survey data and by asking respondents to choose among a list of airlines. The model predicts the airline choice as a function of the socioeconomic characteristics of the passenger as well as the airline service attributes. Nako (1992) studies business travelers' choice of airlines as a function of the airlines' frequent flyer program. He concludes that frequent flyer programs have a positive effect on airline demand and that effect is more significant when the air carrier has a heavy presence at the travelers' home city. Ghobrial and Soliman (1992) estimate a multinomial logit model to estimate the airline city-pair share as a function of peak and off-peak flight frequency, fares, itinerary quality (number of stops), hub presence, and aircraft size. They use data from the top 100 markets in the domestic US markets to estimate the models.

Proussaloglou and Koppelman (1995) use revealed preference survey data to develop multinomial logit models for the air carrier choice between a city pair for business and leisure trips. They use schedule convenience, fares, reliability, market presence, city-pair presence, and frequent flyer program membership as independent variables in these models. Yoo and Ashford (1996) develop logit choice models for international carrier choice behavior in Korea using the trip time, fares, flight frequency, and nationality of the airline as independent variables. They use both the revealed preference data and supplied stated preference data for these models.

Mehndiratta (1996) studies the impact of time-of-day preferences on the scheduling of business trips in the domestic US markets focusing mainly on trips involving air transportation. Mehndiratta divides a regular 24-hour schedule into three periods: work, leisure, and sleep time. He uses a mixed logit model specification to study the impact of disruption of work, leisure, and sleep time on the choice of intercity travel alternatives. The conclusion of this study is that business travelers tend to attribute a higher value to sleep time than to work and leisure time. In addition, the sleep and leisure time spent at home or around home is more valuable than the leisure time and sleep time at a business destination. As a result, travelers avoid staying overnight at their destination unless staying at home and leaving very early in the morning would disrupt their normal sleep schedule too much. Algers and Beser (2001) develop a multinomial logit model for the choice of flight and booking class using bookings data and stated preference data supplied by surveys. The independent variables considered in the model include fares, schedule delay, and booking restrictions.

Proussaloglou and Koppelman (1999) use a logit model to investigate the choice of air carrier, flight, and fare-class. The estimation of these logit models is based on stated preference data. Initial data concerning passenger characteristics (past trips, trip purpose, permanent address, frequent flyer membership) were collected through a mail survey. Then, a random sample of mail survey respondents was

chosen for a phone-based survey designed to simulate the individual traveler's search for information about air travel options and the selection among available alternatives, like during a booking process. The results of the model suggest significant differences in travel behavior between leisure and business travelers.

Coldren and Koppelman (2003 and 2005) extend the itinerary choice models presented in this chapter to provide a representation of the substitution pattern among itineraries. They argue that multinomial logit models do not capture the competition among itineraries that are close in departure time, level-of-service, carrier, or any combination of these dimensions. They account for different substitution patterns between alternatives by developing several generalized extreme value (GEV) models that allow for the possibility of correlation between the error terms for pairs of alternatives.

Adler et al. (2005) use a 2003 Internet-based revealed-preference and stated-preference survey that collects detailed information on the most recent paid domestic air trip of about 600 individuals. The objective of the study is to better understand the trade-offs among the different service characteristics. Several attributes characterizing the itineraries in the stated choice experiments are considered. They include the airline carrier, the airport, the access and egress time, flight times, connections, the fare, the time difference between the desired arrival time at the destination and the scheduled arrival time of the itinerary, the aircraft type, and the on-time performance. The authors use a mixed multinomial logit model to capture the demand sensitivity to the service attributes mentioned earlier. Warbug et al. (2006) estimate standard and mixed multinomial logit models of itinerary choice for business travel based on a stated-preference survey that was conducted in 2001. The results suggest that gender and income level have the most effects on sensitivity to service attributes in itinerary choice behavior. Frequent flyer membership, employment status, travel frequency, and group travel are also considered important determinants.

References

Adler, T., Falzarano, C.S., and Spitz, G. 2005. Modeling Trade-offs in Air Itinerary Choices. *Transportation Research Record*, 1915, 20-26.

Algers, S., and Beser, M. 2001. Modelling Choice of Flight and Booking Class—A Study Using Stated Preference and Revealed Preference Data. *International Journal of Services Technology and Management*, 2(1-2), 28-45.

Ben-Akiva, M.E., and Lerman, S.R. 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand.* The MIT Press. Cambridge, MA.

Coldren, G., and Koppelman, F.S. 2003. Modeling the Competition Among Air Travel Itinerary Shares: GEV Model Development. In 10th International Conference on Travel Behaviour Research, Lucerne, 10-15 August 2003.

Coldren, G., and Koppelman, F.S. 2005. Modeling the Competition among Air-travel Itinerary Shares: GEV Model Development. *Transportation Research, Part A*, 39(4), 345-365.

Coldren, G., Koppelman, F.S., Kasturirangan, K., and Mukherjee, A. 2003. Modeling Aggregate Air Travel Itinerary Shares: Logit Model Development at a Major US Airline. *Journal of Air Transport Management*, 9(6), 361-369.

Ghobrial, A., and Soliman, S.Y. 1992. An Assessment of Some Factors Influencing the Competitive Strategies of Airlines in Domestic Markets. *International Journal of Transport Economics*, 19(3), 247-258.

Koppelman, F.S., and Sethi, V. 2000. Closed-Form Discrete-Choice Models, In D.A. Hensher, and K.J. Button (eds.), *Handbook of Transport Modeling*, Elsevier Science Ltd, Amsterdam.

McFadden, D. 1978. Modeling the Choice of Residential Location. *Transportation Research Record*, 672, 72-77.

Mehndiratta S.R. 1996. *Time of the Day Effects in Inter-City Business Travel.* Ph.D thesis, University of California Berkeley, Berkeley CA.

Nako, S.M. 1992. Frequent Flyer Programs and Business Travellers: An Empirical Investigation. *The Logistics and Transportation Review*, 28(4), 395-414.

Nason, S.D. 1981. The Airline Preference Problem: An Application of Disaggregate Logit. Presented at the 1981 AGIFORS Symposium. Santa Barbara, CA.

Proussaloglou K.E., and Koppelman F.S. 1995. Air Carrier Demand: An Analysis of Market Share Determinants. *Transportation*, 22, 371-388.

Proussaloglou K.E., and Koppelman F.S. 1999. The Choice of Carrier, Flight and Fare Class. *Journal of Air Transport Management*, 5(4), 193-201.

Warburg, V., Bhat, C., and Adler, T. 2006. Modeling Demographic and Unobserved Heterogeneity in Air Passengers' Sensitivity to Service Attributes in Itinerary Choice. *Transportation Research Record*, 1951, 7-16.

Yoo, K.E., and Ashford, N. 1996. Carrier Choices of Air Passengers in PacificRim: Using Comparative Analysis and Complementary Interpretation of Revealed Preference and Stated Preference Data. *Transportation Research Record*, 1562, 1-7.

This page has been left blank intentionally

Chapter 3
Passenger Demand Modeling and Forecasting

Introduction

An important phase of airline planning is to forecast and estimate with acceptable accuracy the number of passengers and the profitability for each flight in the proposed schedule. The increasing sophistication of the airline business has required advanced modeling techniques that are capable of evaluating competition and predicting passenger demand in the different markets. These modeling techniques should provide a detailed representation of: 1) the air carriers' service capacity at the flight, itinerary, and network levels; 2) the travelers' itineraries choice decisions as a function of the characteristics of the available services; and 3) the business processes that might impact on air carriers' competition such as RM methods, code-share agreements, and ticket distribution techniques.

To explain passenger demand forecasting at the flight level, consider two hypothetical airlines competing in the domestic US market, as shown in Figure 3.1. Airline 1 has a hub at Denver, Colorado (DEN) and airline 2 has a hub at Dallas-Fort Worth (DFW), Texas. Consider the case when airline 1 offers a flight from Denver (DEN) to Seattle (SEA) and chooses to evaluate the profitability of this flight. This airline can expect passengers on the DEN-SEA flight to be a combination of local traffic traveling from Denver to Seattle and traffic from other cities to Seattle that connect through Denver. This connecting traffic may, for example, originate from Atlanta (ATL), Chicago (ORD), New York (JFK), Newark (EWR), or Miami (MIA). Therefore, the proposed flight DEN-SEA could be part of several itineraries including DEN-SEA, ATL-DEN-SEA, MIA-DEN-SEA, JFK-DEN-SEA, EWR-DEN-SEA, and ORD-DEN-SEA. It is expected that the number of passengers on the DEN-SEA flight is proportional to the number of passengers selecting these itineraries for their travel.

The expected number of passengers on the flight DEN-SEA depends on a collective set of factors as follows. First, it depends on the size of the markets in which this flight is serving. For example, as the number of passengers traveling to Seattle from Denver, Atlanta, Miami, New York, and Chicago increases, the number of passengers on the flight DEN-SEA is also expected to increase. In addition, the schedule of the DEN-SEA flight affects its number of passengers. The schedule of the flight is pertinent to its departure and arrival time, and convenience of connection with other feeding flights (for example, MIA-DEN, ATL-DEN, and ORD-DEN). Also, the type and size of aircraft of the flight sets an upper limit on the total number of passengers on the flight. Similarly, the size of aircraft of the feeding flights also limits the number of connecting passengers. Furthermore, factors such as marketing,

promotions, ticket distribution channels, and airline reputation affect the number of passengers. Finally, competition with other airlines in each market in which this flight is serving also affects the number of passengers on the flight. For example, in the airline network shown in Figure 3.1, passengers traveling from Miami (MIA) to Seattle (SEA) have two options of travel: connecting with airline 1 at Denver (DEN) or connecting with airline 2 at Dallas-Fort Worth (DFW). Therefore, passenger demand traveling from MIA to SEA will be distributed between these two airlines according to the attractiveness of their schedules. The number of passengers allocated to each airline (itinerary) depends on the number of airlines serving in the market and the attractiveness of the itineraries offered by each airline. In the previous chapter, we showed how itinerary choice models can be used to represent the individuals' itinerary choice behavior. Itinerary choice models present how an individual traveler trades off the different characteristics of the itineraries in the market to select the itinerary that has the best overall characteristics. They are also used to calculate the probability of selecting each itinerary among the different itineraries available to the traveler (the choice set).

Objectives of the Demand Forecasting Model

Itinerary choice models are integrated within a modeling framework for airline demand forecasting and competition analysis. The aim of this framework is to assist

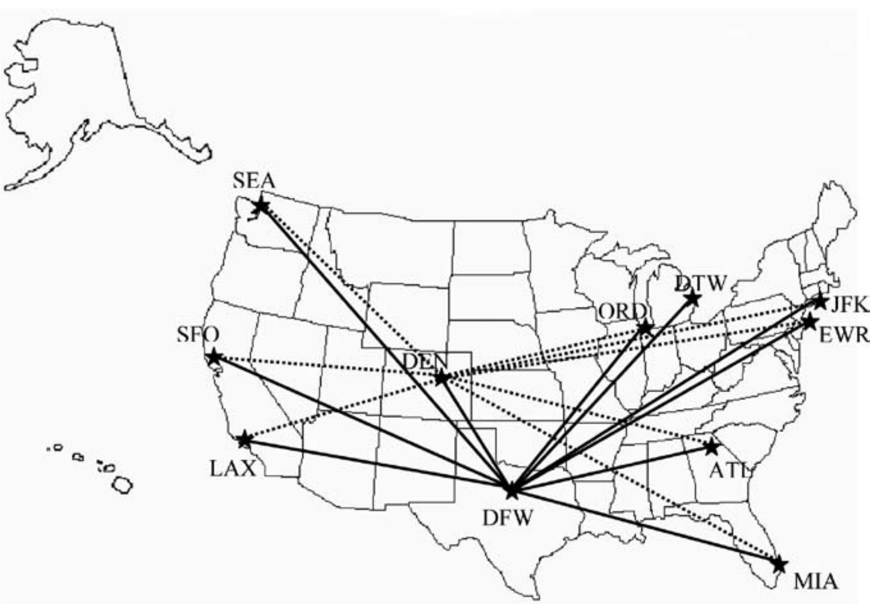


Figure 3.1 Example of network coverage for two competing airlines

airlines to estimate travel demand at the flight level and analyze profitability under different competition scenarios. In particular, the main objectives of passengers' demand models can be summarized as follows:

- To estimate and predict the number of passengers at the flight level.
- To estimate and predict the market share of the airline (the number of passengers that will travel by the airline in a given city pair).
- To perform competition studies and investigate the impact of different operational and strategic decisions (schedule change, pricing, and so on) on passenger demand.
- To perform cost and revenue calculations for the airline schedule.

Data Requirement

Input data required for the model are used to represent an operational scenario of interest. These data include: 1) city-pair travel demand data; 2) flight schedule data; 3) pricing data; and 4) code-share and alliance-related data. Travel demand data are described in terms of the number of travelers expected to travel between all considered city pairs in the network over a pre-specified time period (typically one day). For the US domestic market, these data are widely available through the US Department of Transportation (DOT), Bureau of Transportation Statistics, and several commercial vendors (for example, Data Base Products and OAGBack Aviation Solutions). Flight schedule data are given in terms of the list of scheduled flights for all air carriers. For example, the database of the Official Airline Guide (OAG) (OAGBack Aviation Solutions) provides monthly-updated information on the flight schedule of major scheduled air carriers in the US and worldwide. Each flight in this database is defined through its flight number, origin, destination, departure and arrival times, aircraft type, marketing and operating carriers (flight code share), and day-of-week operation. If a flight is marked as a code-share flight, the list of all air carriers that can market seats on this flight is also given. Pricing data could be fed from the airline's RM and pricing systems. However, most available implementations of the model use the average market fare for each itinerary for all airlines in the market. It should be noted that the model should be capable of having any user-specified input (travel demand, flight schedule scenarios, pricing, and code-share scenarios), which allows studying any operation scenario suggested by the users.

Modeling Framework

The modeling framework presents the interface between the airline demand and supply. This framework adopts a simulation approach in which the airline service capacity and passenger demand are represented at the finest disaggregate level.

Service capacity is tracked at the seat level, and the individual passenger is used for the demand representation. The simulation model replicates the bookings pattern in the different airline reservation systems and estimates the number of passengers on each flight in the schedule. It is capable of simulating a large-scale airline network with multiple competing air carriers.

The overall modeling framework for the airline demand forecasting model is depicted in Figure 3.2. As shown in the figure, given the flight schedule data including code-share agreements, the itinerary-builder module is activated to build all feasible travel itineraries for all city pairs under consideration. This module follows enumeration logic along with a set of predefined constraints to eliminate unfeasible itineraries between each city pair. The main roles for the itinerary builder module can be summarized as follows:

- The origin of the first flight should be the origin of the itinerary.
- The destination of the last flight should be the destination of the itinerary.
- There is a maximum number of connections in the itinerary.
- For connecting itineraries:
  - the origin of a flight should be the destination of its preceding flight
  - the destination of a flight should be the origin of its next flight
  - the departure time of a flight in the itinerary should be greater than the arrival time of its preceding flight
  - the minimum connection time for passengers should be satisfied
  - a maximum connection period should be considered
  - the code sharing of multi-airline itineraries should be considered
  - the itinerary overall travel distance should not exceed a maximum threshold.

The itinerary-building constraints are related to: 1) the itinerary's maximum number of connections; 2) the minimum and maximum connection intervals for passengers; 3) the ratio between the itinerary's total mileage and the direct distance between the city pair; and 4) the multi-airline itineraries in the case of code-share agreements. Also, itineraries with very short connection times (for example, less than 30 minutes) or with very long connection times (for example, more than four hours for domestic trips) are eliminated. Unrealistic connections that generate very long itineraries are also eliminated. For example, an itinerary for a trip between Chicago and San Francisco through Anchorage or Miami is considered infeasible, because it has a considerably long distance compared to the direct travel distance between Chicago and San Francisco. Finally, when there is a code-share agreement between two air carriers on some of their flights, these flights can be used to build multi-airline itineraries. An itinerary that has a code-share flight is represented as many times as the number of the marketing air carriers of this flight.

In an implementation in the US domestic market, a maximum of three connections is considered, which results in almost 3 million itineraries for the US domestic airlines. The output of the itinerary builder module is the set of feasible itineraries

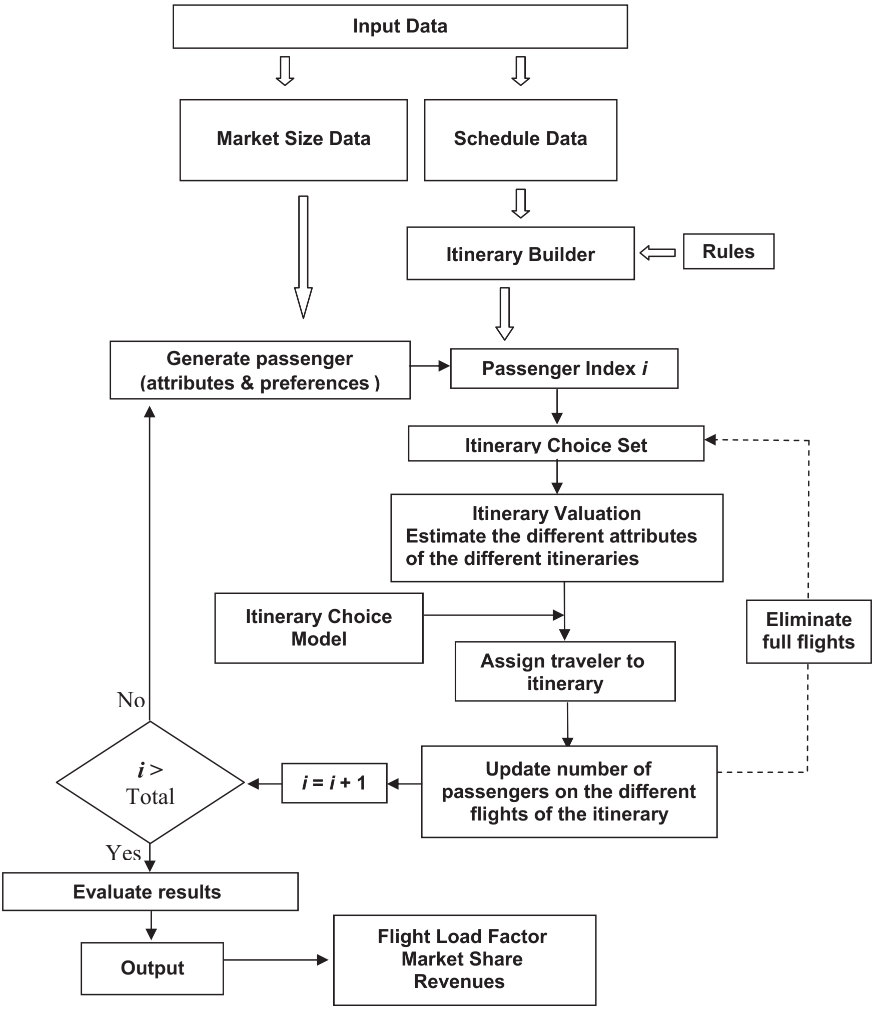


Figure 3.2 Modeling framework for airline competition analysis and demand modeling

for all markets during a given time period (typically one day). Each itinerary is described in terms of its flights along with a set of characteristics, including the number of stops, departure and arrival times, connection intervals, aircraft type for each flight, airfare (price), marketing and operating carriers, and so on.

Given the OD air-travel demand data, passengers are randomly generated. Each passenger is defined by a city pair (market) and a booking date. The booking date of a traveler is generated according to a predefined time distribution specified by the user. Usually, a default built-in uniform distribution is used in case these data are not available. The booking horizon is typically a few months before the flight departure

date. The list of itineraries with available seats is then generated for each traveler, based on their travel market. Travelers evaluate all generated itineraries at the individual level. For this purpose, the discrete choice model developed by Coldren et al. (2003) and presented in the previous chapter is integrated within the simulation model.

Once an itinerary is chosen, the assignment module is activated to update the number of bookings on the chosen itinerary and its corresponding flights. The model tracks the available number of seats for each flight. If the number of bookings reaches the maximum number of seats for sale on the flight, the itineraries that include this flight are eliminated and considered to be closed for sale. Given the new number of bookings on each flight, flight load factors and corresponding market shares are updated accordingly. Also, the total revenue and the profitability of the schedule can be calculated.

Illustrative Example

Consider the hypothetical network as shown in Figure 3.3, where there are three different air carriers operating in this network. The flights of each air carrier are given by different line patterns. Air carrier *1* has a solid pattern; air carrier *2* has a dashed pattern; and air carrier *3* has a dotted pattern. The total daily demand from A to B is 200 travelers; from A to C it is 300; and from A to D it is 20. The main characteristics of the different itineraries of the three air carriers are given in Table 3.1. The following illustration shows how air carrier *1* estimates the number of passengers on flight AD.

First, it should be noted that flight AD offered by air carrier *1* is part of the itineraries in three different markets, which include the nonstop itinerary from A to D, the single-connect itinerary from A to B, and the single-connect itinerary from A to C. Passengers traveling in these markets could be using flight AD for their travel. Second, we investigate the air carrier market share and level of competition in each of these markets.

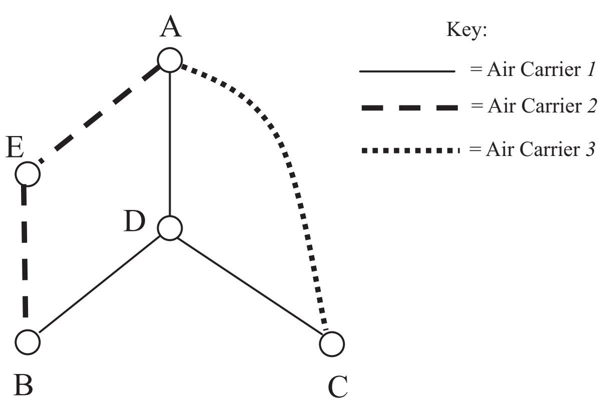


Figure 3.3 Example of a hypothetical network of three air carriers

Table 3.1 The main characteristics of the different itineraries of the three air carriers

Variables	Airline 2	Airline 1		Airline 3
		AB	AC	
Level-of-Service				
Nonstop Itinerary in Nonstop Market				1
Direct Itinerary in Nonstop Market				
Single-Connect Itinerary in Nonstop Market			1	
Double-Connect Itinerary in Nonstop Market				
Direct Itinerary in Direct Market				
Single-Connect Itinerary in Direct Market				
Double-Connect Itinerary in Direct Market				
Single-Connect Itinerary in Single-Connect Market	1	1		
Double-Connect Itinerary in Single-Connect Market				
Connection Quality				
Second-Best Connection				
Second-Best Connection Time Difference				
Best Connection Time Difference	12			
Distance Ratio	105	100	119	100
Carrier Attributes				
Point of Sale Weighted City Presence	30	15	20	20
Fare Ratio	110	95	103	120
Code Share				
Aircraft Size and Type				
Propeller				
Regional Jet				
Propeller Seats				
Regional Jet Seats				
Mainline Seats	200	150	170	170

Table 3.1 ***Concluded***

Variables	Airline 2	Airline 1		Airline 3
		AB	AC	
Time of Day				
Midnight–5 AM				
5–6 AM				
6–7 AM				
7–8 AM				
8–9 AM	1			
9–10 AM				
10–11 AM				
11–12 noon				
12–1 PM				
1–2 PM		1		
2–3 PM				
3–4 PM				1
4 –5 PM				
5–6 PM			1	
6–7 PM				
7–8 PM				
8–9 PM				
9–10 PM				
10–Midnight				

Market AD

Air carrier *1* is the only air carrier working in market AD; accordingly, its market share is 100 percent. All passengers flying from A to D select the air carrier *1* for their travel. This selection counts for 20 passengers on flight AD.

Market AB

There are two air carriers competing in the market AB (air carrier *1* and air carrier *2*) with two different itineraries. Air carrier *1* has a single-connect itinerary that is connecting at point D, and air carrier *2* has a single-connect itinerary that is

connecting at point E. The itinerary choice models presented in the previous chapter are used to estimate the probability of a traveler selecting each of these two itineraries. Assuming the market AB is an E-E market, the probability of selecting air carrier *1* is estimated to be 0.467, and the probability of selecting air carrier *2* is 0.533. Accordingly, on average, out of the 200 passengers flying from A to B, about 93 passengers (0.467 × 200) select air carrier *1* and use flight AD.

Market AC

There are two air carriers competing in the market AB (air carrier *1* and air carrier *3*), with two different itineraries. Air carrier *1* has a single-connect itinerary that is connecting at point D, and air carrier *3* has a nonstop itinerary. Again, the itinerary choice models are used to estimate the probability of selecting each of these two itineraries. Assuming the market AC is an E-E market, the probability of selecting air carrier *1* is estimated to be 0.039, and the probability of selecting air carrier *3* is 0.961. Accordingly, on average, out of the 300 passengers flying from A to C, about 12 passengers (0.039 x 300) select air carrier *1* and use flight AD.

Accordingly, in total, the total number of passengers estimated for flight AD = 20 + 93 + 12 = 125 passengers.

Market Share and the Herfindahl-Hirschman Index (*HHI*)

An important output of the air carrier demand forecasting model is the air carrier share of passengers in each market, which is known as the Quantitative Share Index (*QSI*). The *QSI* for a particular air carrier in a given market is defined as the percentage of passengers who are using this air carrier in this market. It is calculated as follows for air carrier *i* and market *j*:

$$QSI_i^j = \frac{N_i^j}{\sum_{i \in I} N_i^j} \tag{3.1}$$

Where:

N_i^j = number of passengers flown by air carrier *i* in market *j*

I = total number of air carriers operating in market *j*.

$$0\% \le QSI_i^j \le 100\% \tag{3.2}$$

Air carriers keep track of their market share record to make sure that it is consistent with the service they offer and with the level of competition in the market.

Another index that is used to measure the level of competition among service providers is the Herfindahl-Hirschman Index (*HHI*), a widely accepted measure of market concentration or market competitiveness. It is calculated by squaring

the market share (*QSI*) of each service provider in the market and summing the resulting numbers. The *HHI* is calculated for market *j* as follows:

$$HHI^{j} = \sum_{i \in I} (QSI_i^j)^2 \tag{3.3}$$

The *HHI* increases as the number of service providers (air carriers) in the market decreases and as the disparity in size between those providers increases. The increase in the *HHI* value indicates a high concentration in the market (that is, few providers control the market). To explain how the *HHI* is used, consider two hypothetical markets. The first market has two air carriers providing service with fair competition, where each air carrier has a 50 percent market share. Therefore, the *HHI* for this market is equal to:

$$HHI^{1} = 0.5^2 + 0.5^2 = 0.5$$

The second market has two air carriers that provide service, where one air carrier controls the majority of the market. The first air carrier has a market share (*QSI*) of 90 percent, and the second air carrier has a market share (*QSI*) of 10 percent. The *HHI* for this market is equal to:

$$HHI^{2} = 0.9^2 + 0.1^2 = 0.82$$

It is clear that the market that has a high market concentration also has a higher *HHI*. Theoretically, the *HHI* ranges between zero and one. The *HHI* is equal to zero when there is an infinitely large number of service providers serving in the market, where each one has a small market share (≈ 0). The *HHI* is equal to one when there is only one service provider serving in the market, with a market share of 100 percent (pure monopoly).

Primary Contributions

Several modeling frameworks have been developed to model the air-travel demand at the network level. These models vary in their objective and the level of detail of their representation of the different components, including demand, network capacity, and interaction. Ghobrial and Kanafani (1995) present a network-equilibrium model to investigate network structure and operation policies. The model is applied to predict the future of hubbing and dehubbing in the US airport system. Sabre® (www.sabresolutions.com) has developed a model called AirFlite™ to evaluate flight profitability for airlines. The model, which is used by several major airlines in the US, adopts several alternatives of assignment models. These models include discrete choice models that estimate the market share of the different itineraries in the different OD markets. Based on these estimated market shares, the model assigns passengers in the different OD pairs to the different itineraries.

The assignment is performed at the macro level for the whole demand in the OD pair, which is considered as one of the main limitations of these models. Therefore, the model needs several iterations of spill and recapture to adjust passenger counts based on the available seat capacity. This macro assignment precludes capturing the impact of decisions that are preformed at the individual level, such as the response to airfare differences (RM strategies) and seat availability.

A simulation model called the Passenger Origin-Destination Simulator (PODS) has been developed by researchers at the Massachusetts Institute of Technology (MIT) and a consortium of seven leading air carriers in the US The main goal of this model is to examine the impact of RM methods, especially seat allocation decisions at the network level (Zickus 1998, Darot 2001, and Carrier 2003). The model classifies prospective travelers into two main categories; leisure and business. Available travel options are generated separately for each group based on their preferences and the time windows of their travel. A discrete choice mixed logit model has recently been added to the PODS's framework to capture how passengers evaluate the different available itineraries (Carrier 2003).

Frank et al. (2006) present an event-driven stochastic simulation model to replicate a real-world RM system. The main objective of this model is to evaluate the impact of fleet assignment adjustments on revenue during the booking horizon. A special demand model is developed to consider dependencies between booking classes for the simulation.

Abdelghany and Abdelghany (2008) present a modeling framework for Airlines Competition Analysis and Demand Modeling (ACADM). The framework adopts a micro-simulation modeling that replicates how prospective travelers select from travel options. The model simulates traveler itinerary choices at the disaggregate level and provides several performance metrics at the flight, market, and system levels. Several simulation experiments are conducted using data from the US domestic markets with 124 airlines and 17,500 flights. Abdelghany and Abdelghany (2008) use ACADM to examine the trade-offs between two common types of ticket distribution channel: 1) distribution channels with high market penetration and high competition among subscribed airlines (for example, Travelocity, Expedia, and Orbitz) versus 2) distribution channels with low market penetration and low airline competition (for example, aa.com and continental.com). The results provide useful insights for airlines and other stakeholders of ticket distribution strategies.

References

Abdelghany, A., and Abdelghany, K. 2008. A Micro-Simulation Approach for Airlines Competition Analysis and Demand Modeling. *International Journal of Revenue Management*, 2(3), 287-306.

Carrier, E. 2003. Modeling Airline Passenger Choice: Passenger Preference for Schedule in the Passenger Origin-Destination Simulator (PODS). M.S. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Coldren, G., Koppelman, F.S., Kasturirangan, K., and Mukherjee, A. 2003. Modeling Aggregate Air Travel Itinerary Shares: Logit Model Development at a Major U.S. Air Carrier. *Journal of Air Transport Management*, 9(6), 361-369.

Darot J. 2001. Revenue Management for Airline Alliances: Passenger Origin-Destination Simulation Analysis. M.S. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Frank, M., Friedemann, M., Mederer, M., and Schroeder, A. 2006. Airline Revenue Management: A Simulation of Dynamic Capacity Management. *Journal of Revenue and Pricing Management*, 5(1), 62-71.

Ghobrial, A., and Kanafani, A. 1995. Future of Airline Hubbed Networks: Some Policy Implications. *Journal of Transportation Engineering*, 121(2), 124-134.

Zickus, J.S. 1998. Forecasting for Airline Network Revenue Management; Revenue and Competitive Impacts. M.S. thesis, Massachusetts Institute of Technology, Cambridge, MA.

SECTION II
Scheduling of Resources

This page has been left blank intentionally

Chapter 4
Fleet Assignment

Introduction

Most major air carriers operate different aircraft types (fleets), which give them the flexibility to compete in markets with different characteristics, such as demand and length of haul. For example, Table 4.1 gives the fleet mix operated by most major US carriers (as well as confirmed aircraft orders), as of Summer 2008. As shown in the table, there is no single air carrier that operates one fleet type, and most air carriers operate about ten different fleets. Thus, air carriers must find the optimal allocation of flights to their different fleet types: a problem known as the fleet assignment problem. This process determines the assignment of an aircraft type to each flight in the schedule (timetable). The main objective of fleet assignment is to match the characteristics of the aircraft type and the flight to minimize the total cost to the air carrier.

Two main characteristics of the aircraft are considered when assigned to a particular flight. First, the aircraft's seat capacity should be comparable to the expected passenger demand on the flight. Second, the distance between the origin and the destination of the flight should be within the aircraft's flight range. The maximal total range is the distance an aircraft can fly between takeoff and landing, as limited by fuel capacity. The fuel limit for an aircraft is fixed by the fuel load and rate of consumption. The range of the aircraft also affects its ability to fly above water, such as across the Pacific or Atlantic Ocean. Table 4.2 gives the values of these two characteristics for the most common aircraft fleets used by US carriers.

Other fleet characteristics include the economics of fuel consumption. Each aircraft type has an optimal range of travel distance in which it produces the best fuel consumption performance. Also, large aircraft might not be able to land at small airports due to runway weight limitations or gate configuration. In addition, most airports charge different landing fees based on the aircraft's weight or capacity. Another factor that might affect the operation of an aircraft type at an airport is noise level. Aircraft with high noise levels may not be allowed to land at some airports or might be subject to curfews. In addition, the fleet maintenance requirements might have an impact on the flights that are assigned to the different fleets. For example, the location of the fleet maintenance station might require assigning the fleet to one or more of the flights that start or end at the maintenance station.

Finally, there are logical constraints that should be satisfied in the fleet assignment problem. For instance, at any moment of time, the number of flights that are assigned to any fleet type should not exceed the number of aircraft in this

Table 4.1 Aircraft fleets for major US air carriers (Summer 2008)

Fleet	American Airlines		Southwest Airlines		United Airlines		Delta Air Lines		Continental Airlines	
	aircraft	seat	aircraft	seat	aircraft	seat	aircraft	seat	aircraft	seat
McDonnell Douglas DC-9-30										
McDonnell Douglas DC-9-40										
McDonnell Douglas DC-9-50										
McDonnell Douglas MD-82	243	136–140								
McDonnell Douglas MD-83	93	136–140								
McDonnell Douglas MD-90							16	150		
McDonnell Douglas MD-88							117	134–142		
Boeing 717-200										
Boeing 777-200					19	258–348				
Boeing 777-200LR							2–6 orders	276		
Boeing 777-200ER	49–5 orders	245–247			33	253–269	8	268	20–8 orders	283–285
Boeing 767-200ER	16	165–167							10	174
Boeing 767-400ER							21	246–285	16	235–256
Boeing 767-300ER	58	219–225					59	216–217		
Boeing 767-300					35	183–224	21	250–262		
Boeing 757-200	125	188			97	110–182	135	174–184	41	175

fleet. Another logical constraint is to maintain the continuity of fleet types at the different airport stations. In other words, at any station, the number of inbound flights assigned to a particular fleet type should equal the number of outbound flights that are assigned to this fleet type.

Graphical Representation of the Fleet Assignment Problem

The main input to the fleet assignment problem is the timetable of the air carrier, which represents the schedule of the proposed flights. Each flight is defined by the origin, destination, and scheduled departure time. The expected passenger (and possibly cargo) demand for each flight is also given. On the supply side, fleet

Table 4.1 *Continued*

Fleet	Northwest Airlines		US Airways		JetBlue Airways		AirTran Airways		Alaska Airlines	
	aircraft	seat	aircraft	seat	aircraft	seat	aircraft	seat	aircraft	seat
McDonnell Douglas DC-9-30	42	100								
McDonnell Douglas DC-9-40	11	110								
McDonnell Douglas DC-9-50	34	125								
McDonnell Douglas MD-82										
McDonnell Douglas MD-83									7	140
McDonnell Douglas MD-90										
McDonnell Douglas MD-88										
Boeing 717-200							87	117		
Boeing 777-200										
Boeing 777-200LR										
Boeing 777-200ER										
Boeing 767-200ER			10	203						
Boeing 767-400ER										
Boeing 767-300ER										
Boeing 767-300										
Boeing 757-200	55	160–184	44	179–193						

information including count of aircraft and the main characteristics are also given. The main objective of the fleet assignment problem is the assignment of an aircraft type to each flight in the schedule (timetable).

Time-staged Flight Network

Before proceeding with the main formulation of the fleet assignment problem, consider the graphical representation of the air carrier timetable which is known as the time-staged flight network. Figure 4.1 shows a representation of part of a hypothetical air carrier schedule. As shown in the figure, the different airports are represented by vertical lines. Time is represented vertically along the vertical line of each station. Each flight is represented by a diagonal arc (arrow) that connects

Table 4.1 Continued

Fleet	American Airlines		Southwest Airlines		United Airlines		Delta Air Lines		Continental Airlines	
	aircraft	seat	aircraft	seat	aircraft	seat	aircraft	seat	aircraft	seat
Boeing 757-300			188	137					17	216
Boeing 737-800	77–29 orders	142–148					71–34 orders	150–160	111–6 orders	150–157
Boeing 737-300					64	120–128			47	124
Boeing 737-500			25	122	30	104–110				
Boeing 737-700			320–121 orders	137			10 orders	124	36–44 orders	124
Boeing 737-400										
Boeing 737-400F										
Boeing 737-400C										
Boeing 737-900									12	167–169
Boeing 737-900ER									10–28 orders	173
Boeing 747-400					30	347–374				
Boeing 787-8									8 orders	
Boeing 787-9									17 orders	
Airbus A319					55	120				
Airbus A320					97	138–156				
Airbus A321										
Airbus A330-200										
Airbus A330-300										
Airbus A350-800XWB										
Airbus A350-900XWB										
Airbus A300B4-600R	34	267								
Embraer 190										

Table 4.1 *Concluded*

Fleet	Northwest Airlines		US Airways		JetBlue Airways		AirTran Airways		Alaska Airlines	
	aircraft	seat	aircraft	seat	aircraft	seat	aircraft	seat	aircraft	seat
Boeing 757-300	16	224	46	126–134						
Boeing 737-800									36	157 (16/141)
Boeing 737-300										
Boeing 737-500										
Boeing 737-700							54–65 orders	137	20	124
Boeing 737-400			40	144					34	132
Boeing 737-400F									1	
Boeing 737-400C									5	72
Boeing 737-900									12	172
Boeing 737-900ER										
Boeing 747-400	16	403								
Boeing 787-8	18 orders	200								
Boeing 787-9										
Airbus A319	57–5 orders	124	98–4 orders	124						
Airbus A320	73–2 orders	148	77–49 orders	150	107–70 orders	150				
Airbus A321			28–37 orders	183						
Airbus A330-200	11	243	25 orders	264						
Airbus A330-300	21	298	9	293						
Airbus A350-800XWB			18 orders							
Airbus A350-900XWB			4 orders							
Airbus A300B4-600R										
Embraer 190			19–6 orders	99	34–67 orders	100				

Table 4.2 Flight range and seat capacity for most common aircraft fleets

		Flight Range (Km)									
		5,600 to 5,900	6,800 to 7,700	9,000 to 10,000	10,500 to 11,300	12,250 to 12,500	13,300 to 13,900	14,300 to 14,500	15,000 to 15,200	15,650 to 16,000	16,700 to 17,400
Seat Capacity	90–132	737–600									
	110–130 or 120–149	A318, 737–700, A320	A319	737–700ER							
	162–189	737–800									
	181–255	A321, 737–900	757–200	A310–200, A310–300	767–300ER	767–200ER		A340–200	787–8		
	243–375		757–300		767–400ER, Boeing 747SP						
	253–293	A300	A300–600			A330–200				A350–800, 787–9	
	295–335	787–3			A330–300				A350–900		
	313–366									A340–500	A340–500HGW, A350–900R
	350							A350–1000			
	295–440						A340–300, B747–400	777–200ER, B747–400ER			777–200LR
	358–550		747–100SR, B747–300SR	747–100		747–200		777–300ER			
	380–419							A340–600		A340–600HGW	
	467							747–8			
	525–853								A380		

between two airports. The start and end of the diagonal arrow represents the departure and arrival times of the flight, respectively.

The time-staged flight network is extended to represent the case when aircraft are assigned to the different flights, as shown in Figure 4.2. This extended network has three different types of arc: the flight arc, ground arc, and remaining overnight (RON) arc. A flight arc in the network corresponds to the movement of an aircraft of type *e* along a flight leg *f*. The ground arc represents the time at which the aircraft at any station connects between an inbound and an outbound flight. The ground time is required for deplaning passengers on the inbound flight, unloading luggage and cargo, aircraft cleaning, refueling, boarding passengers on the outbound flight, and loading luggage and cargo. The ground time depends on the aircraft type; a large aircraft usually requires longer ground periods. Finally, the RON arc corresponds to the case when the aircraft stays overnight at the station to operate an outbound flight on the next day.

At each station there is a set of inbound flights (arrivals) and another set of outbound flights (departures) that have different characteristics, including expected demand and haul length. These flights are assigned to different aircraft types. For each aircraft type at the station, the number of inbound flights assigned to this aircraft should equal the number of outbound flights assigned to it. This condition

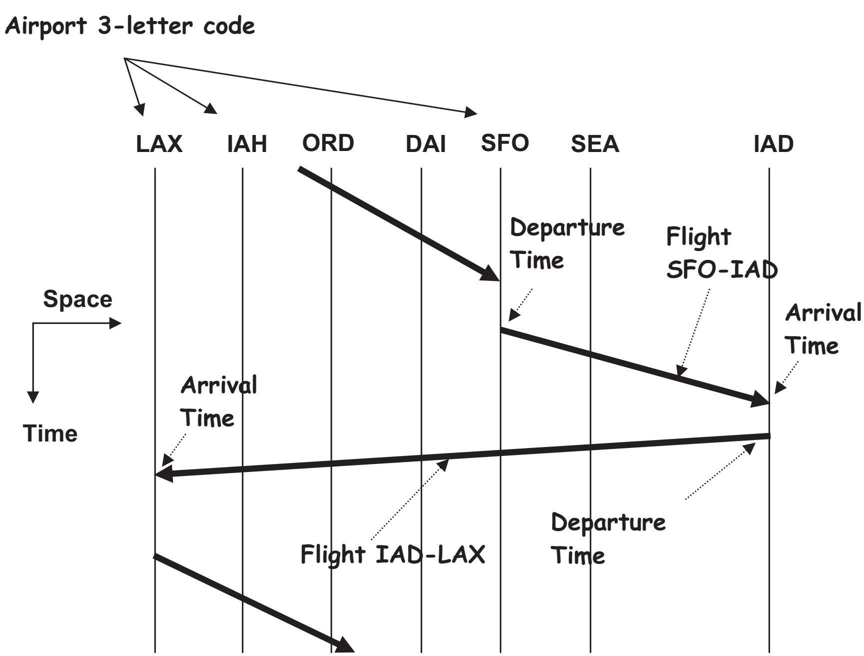


Figure 4.1 Time-staged flight network

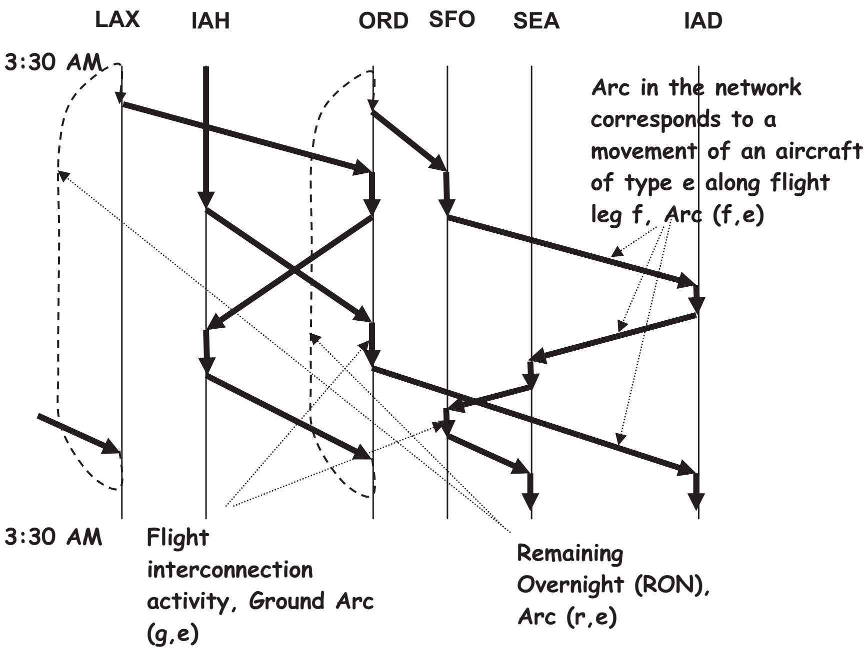


Figure 4.2 Time-staged flight network with aircraft assignment

guarantees the continuity of the aircraft at the different stations. Once an aircraft of a particular type lands at a station serving any inbound flight, a suitable outbound flight for this aircraft is available to take it out of the station. For example, Figure 4.3 shows the inbound and outbound flights at an airport station. Each flight is represented by an arrow. The shape of each arrow represents the aircraft type to which the flight is assigned. As shown in the figure, there are three aircraft types and, for each aircraft type at the station, the number of inbound flights is equal to the number of outbound flights.

Interconnection Nodes

The movements of a particular aircraft type at a given station can be classified into interconnection nodes. The interconnection node consists of the aircraft serving the arrivals and the aircraft serving the following set of departures that occur before the next arrival. The node also includes the aircraft on the ground before the first aircraft arrival in the node and the aircraft on the ground after the last departure of the node. The node also includes the aircraft that do not depart before the next aircraft arrival occurs.

To maintain the continuity condition of this aircraft type, for every node and for every aircraft type, the total number of aircraft into the node should be equal

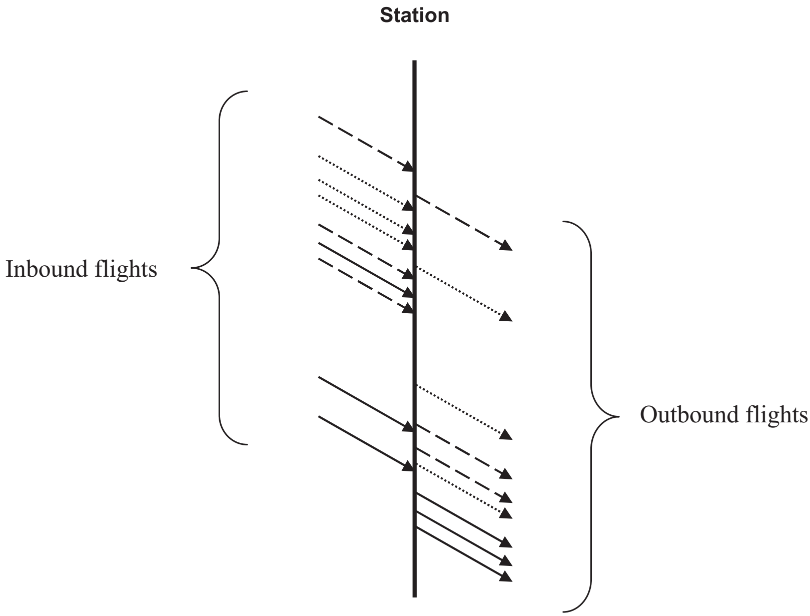


Figure 4.3 Inbound and outbound flights at a hypothetical station for three aircraft types

to the total number of aircraft out from the node. The aircraft into the node are from flight arrivals and aircraft remaining from the previous interconnection node. The aircraft out from the node are aircraft serving departing flights and aircraft that remain on the ground with the next node. Figure 4.4 presents an example of two interconnection nodes for aircraft of type *e* at airport *a*. These nodes are denoted node *n* and node *n* + 1, respectively. For node *n*, *I* [*a, e, n*] and *O* [*a, e, n*] represent the arrivals and departures of aircraft of type *e* at airport *a*. The total aircraft input to node *n* is represented by *IN*[*a, e, n*], and the total output is represented by *OUT*[*a, e, n*]. Another example of interconnection nodes is presented in Figure 4.5, which shows the last two interconnection nodes in the day for aircraft type *e* at airport *a*. For the last node, the aircraft that stay on the ground are the ones that remain overnight (RON) at station *a*. These aircraft represent an input to the first interconnection node of the next day.

The Main Requirements of the Fleet Assignment Solution

This section presents the mathematical formulation of the general fleet assignment problem (Rushmeier et al. 1995). As mentioned earlier, the goal of fleet assignment is to assign aircraft types to a flight leg so as to minimize cost

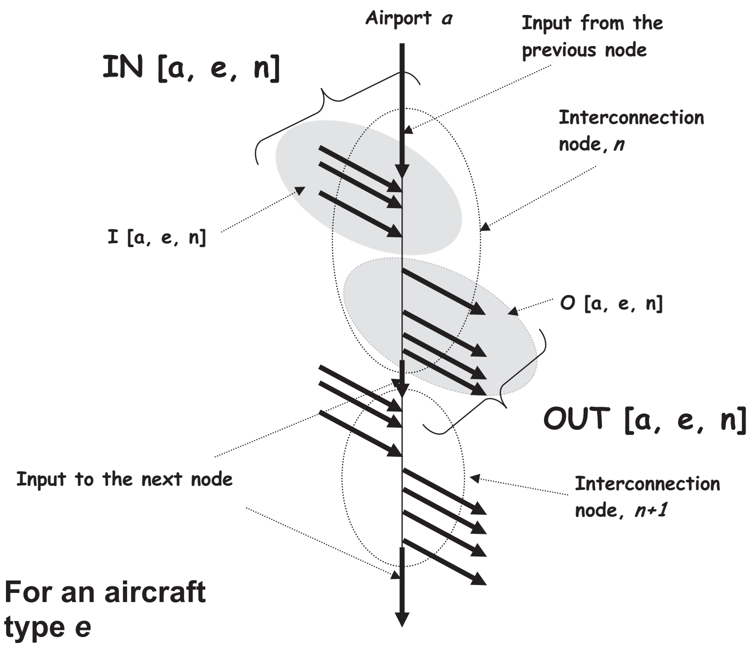


Figure 4.4 Illustration of an interconnection node at a station

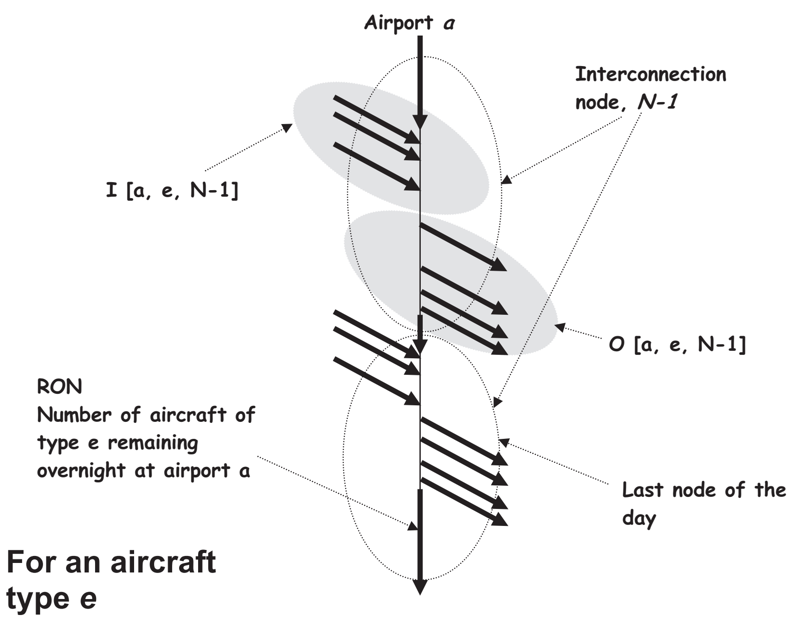


Figure 4.5 Illustration of the last interconnection node at a station

or maximize profitability. For each flight, a decision variable can be defined as follows:

$$x_{fe} = \begin{cases} 1 & \text{if flight leg } f \text{ is assigned to aircraft type } e \\ 0 & \text{otherwise} \end{cases} \tag{4.1}$$

Associated with each value of the decision variable x_{fe} is a cost function or value function, which is denoted by c_{fe}. The cost function is the difference between the operating cost of aircraft type *e* when assigned to flight leg *f* and the revenue generated when aircraft type *e* is assigned to flight leg *f*. The operating cost depends on the fleet type, fuel burn rate, length of the trip, landing fees, crew cost, and so on. The generated revenue is a function of the expected passenger demand to be accommodated by aircraft type *e* and the airfare. It also depends on cargo demand and tariffs, if any.

The objective function is to minimize the total cost of fleet assignment for all the flights in the schedule, which can be represented mathematically as follows:

$$C = Minimize \sum_{f \in F} \sum_{e \in E(f)} c_{fe} \, x_{fe} \tag{4.2}$$

Where:

F = set of flights in the schedule

$E[f]$ = set of aircraft types that can operate flight leg *f*.

The set $E[f]$ is used in the formulation because it is known that some aircraft types cannot fly some flights. For example, noise level, runway requirements, and above-water flying might prevent some aircraft types from operating some flights in the air carrier timetable. The following set of constraints needs to be verified for a feasible fleet assignment solution.

Coverage Constraints

A feasible solution requires that exactly one aircraft type is to be assigned to each flight leg in the schedule. This constraint, which is known as the flight coverage constraint, can be represented mathematically as follows:

$$\sum_{e \in E[f]} x_{fe} = 1 \qquad \forall f \in F \tag{4.3}$$

Given $x_{fe} \in (0,1)$, these constraints ensure that exactly one aircraft type is chosen for flight *f*.

Sizing Constraints

The second set of constraints for the fleet assignment problem is that the total number of used aircraft of any type should not exceed the total available aircraft of this type. Consider p_e as the number of aircraft available for aircraft type e. At any time, the total number of aircraft of type e that are assigned to flights, on the ground connecting between two flights, and remaining overnight at any station should be less than p_e. This constraint, which is known as the sizing constraint, is presented in Figure 4.6. Figure 4.6 presents the flight assignment to aircraft type e. A counting time is designated (for instance, at 5:00 AM), and the number of aircraft, ground, and remaining overnight (RON) arcs that cross the designated counting time is recorded. This number should be less than the available number of aircraft of type e. This set of constraints can be represented mathematically as follows:

$$\sum_{j \in AC_e} x_{je} \leq p_e \qquad \forall e \in E \tag{4.4}$$

Where:
AC_e = set of flight, ground, and overnight arcs of aircraft type e that cross the designated counting time
E = set of aircraft types.

For example, in the fleet assignment solution shown in Figure 4.6 for a particular aircraft type, four aircraft cross the designated counting time, 5:00 AM, as follows: RON at LAX, RON at ORD, a flight out of SEA, and an aircraft on the

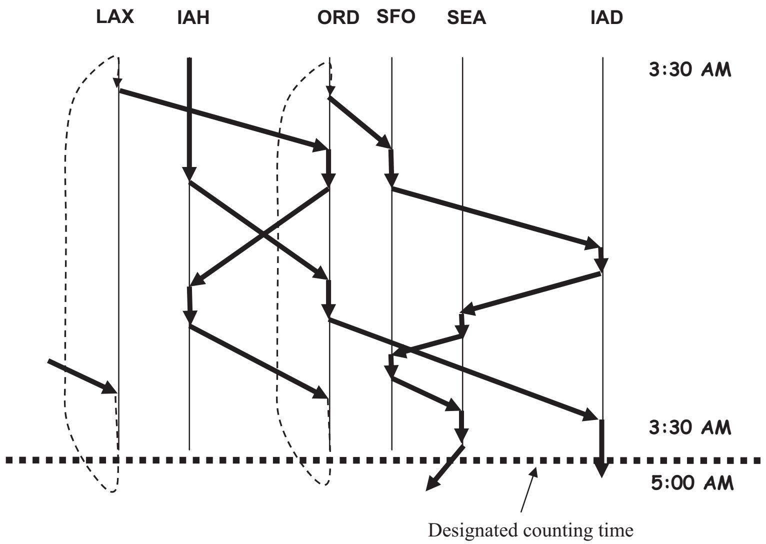


Figure 4.6 Illustration of the sizing constraints

ground at IAD. This solution is feasible if and only if the air carrier has at least four aircraft of this type.

It should be noted that if the given number of flights to be fleeted is high (that is, if too many flights are proposed in the air carrier's schedule), the sizing constraints might give an infeasible solution. To avoid infeasibility, this set of constraints can be written as follows:

$$\sum_{j \in AC_e} x_{je} + u_e - o_e = p_e \qquad \forall e \in E \tag{4.5}$$

Here, u_e and o_e are two variables that represent the under-use and over-use of aircraft type *e*, respectively. Only u_e or o_e is greater than zero. When $u_e > 0$, the aircraft type *e* is under-used. Conversely, when $o_e > 0$, the aircraft type *e* is over-used. In the objective function, a large penalty associated with the aircraft type over-use is considered. Also, a very small benefit is incorporated for under-use of each aircraft type.

Continuity Constraints

The third set of constraints for the fleet assignment problem is to maintain continuity of aircraft types to guarantee that, at each station, any arriving aircraft of any type can leave on a valid departing flight. The continuity condition can also be satisfied at the level of the whole day or at the level of the interconnection node defined above. For every node and for each aircraft type, the total number of aircraft into the node should equal the total number of aircraft out from the node. The aircraft into the node are from flight arrivals and aircraft remaining from the previous interconnection node. The aircraft out from the node are aircraft serving departing flights and aircraft that remain on the ground for the next node. For example, for node *n* in Figure 4.7, the aircraft into the node from flight arrivals is equal to three. Aircraft remaining from the previous interconnection node is equal to two. The aircraft out from the node that are serving the departing flights is equal to four, and the aircraft that remain on ground for the next node is equal to 1. The total aircraft into the node, which is five, is equal to the total aircraft out from the node.

The continuity constraints can be represented mathematically as follows:

$$\sum_{j \in IN(a,e,n)} x_{je} = \sum_{j \in OUT(a,e,n)} x_{je} \qquad \forall e \in E, n \in N, a \in A \tag{4.6}$$

These constraints indicate that for any node *n* at airport *a*, the number of aircraft of a certain type *e* that are inbound to the node should equal the number of outbound aircraft.

Where:

N = set of interconnection nodes

A = set of airports
E = set of aircraft types.

Through-flight Constraints

A through flight is when an aircraft traveling from point A to point B stops at intermediate station C to drop off and pick up passengers. As passengers traveling from point A to point B do not change aircraft at the intermediate station they do not have to connect between two gates. In many cases, air carriers schedule through flights at one or more of their airports to increase the convenience to passengers. The fleet assignment solution should consider the requirement of through flights in the schedule. In particular, the inbound and outbound flights of the through flight should be of the same aircraft type. For example, Figure 4.8 shows a hypothetical schedule of an air carrier at one of the airports. If the air carrier is planning that F1-F2 should be a through flight at the station, then the fleet assignment solution should guarantee that both flights F1 and F2 are assigned to the same equipment type. The through-flight constraint for any two flights F1 and F2 can be represented mathematically as follows:

$$x_{F^1 e} = x_{F^2 e} \tag{4.7}$$

This constraint guarantees that the two flights F1 and F2 are from the same aircraft type $e \in E[\text{F }1] \cap E[\text{F }2]$.

Maintenance Constraints

The fleeting of the schedule must take into consideration the location and timing of maintenance activities. Maintenance constraints take a number of different forms, which together reflect most of the maintenance issues faced by an air carrier.

First, because most aircraft maintenance activities are performed overnight and at specific stations that have maintenance facilities, the fleet assignment problem needs to ensure that a minimum number, m_e, of aircraft of fleet type e remain overnight at a specified group of airports that serve as maintenance stations for that type of aircraft.

These constraints can be represented mathematically as follows:

$$\sum_{r \in RON_{[r]}} x_{re} \geq m_e \quad \forall e \in E \tag{4.8}$$

Where:
$RON_{[e]}$ = set of RON arcs for aircraft type e at the stations that serve as maintenance stations for aircraft type e.

For example, consider the fleet assignment solution considered in Figure 4.6 for aircraft type e. Assume that this air carrier locates its maintenance facilities at LAX and ORD. Assume also that the minimum number of aircraft from fleet type

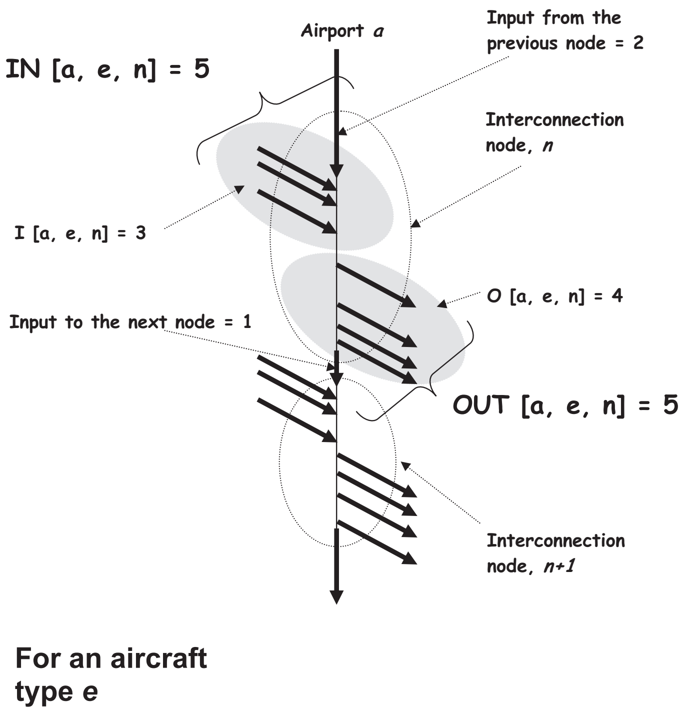


Figure 4.7 Illustration of the aircraft count at the interconnection node

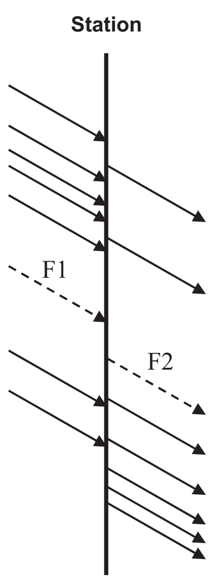


Figure 4.8 Example of a through flight

e required to remain overnight at LAX and ORD is equal to two (that is, $m_e = 2$). The solution of Figure 4.6 satisfies the maintenance constraints of two aircraft remaining overnight at LAX and ORD, so maintenance activities can be performed for these two aircraft. If it is required that three aircraft remain overnight at both LAX and ORD (that is, $m_e = 3$), a solution similar to the one shown in Figure 4.9 may be obtained. In this solution, two aircraft remain overnight at ORD, and one aircraft remains overnight at LAX.

This maintenance constraint might result in a solution that has aircraft remaining overnight at the maintenance stations in a way that is not proportional to the capacity of each maintenance station. For example, the solution in Figure 4.9 allocates two aircraft to remain overnight at ORD and one aircraft at LAX. This could be an inappropriate solution if ORD can handle the maintenance activity of one aircraft only, while LAX can handle two aircraft. To overcome this problem, a second set of maintenance constraints should be added to specify the minimum and maximum number of aircraft that may remain overnight at the maintenance station. These two limits are comparable to the maintenance capacity of the station.

This set of constraints can be represented mathematically as follows:

$$l_e^s \geq \sum_{r \in RON_{[e,s]}} x_{re} \geq u_e^s \quad \forall e \in E, \forall s \in S \tag{4.9}$$

Where:

$RON_{[e,s]}$ = set of RON arcs for aircraft type *e* at maintenance station *s* that serves as maintenance station for aircraft type *e*

l_e^s = minimum number of aircraft of type *e* that may remain overnight at maintenance station *s*

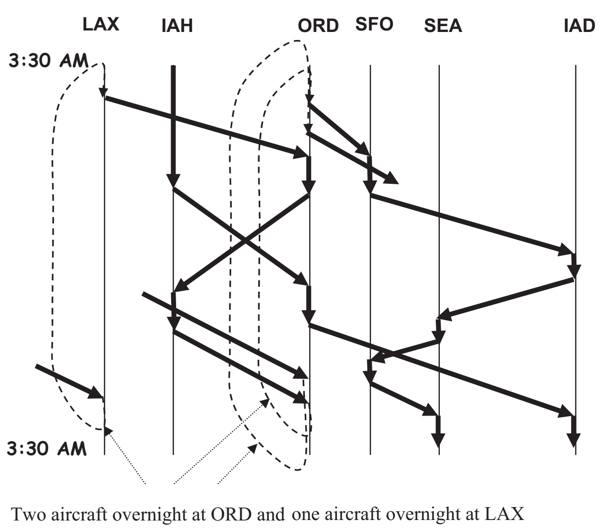


Figure 4.9 Example of three aircraft overnight at ORD and LAX

u_e^s = maximum number of aircraft of type e that may remain overnight at maintenance station s.

In some cases, an aircraft is required to remain for a minimum time period at a particular maintenance station, possibly for a special maintenance activity that requires a longer time and is only performed at this station. Therefore, the RON period of the aircraft at this station should be greater than the minimum required maintenance time. To explain this constraint, which is known as the time-window constraint, consider the arrivals and departures of a particular aircraft type at one of the maintenance stations, as presented in Figure 4.10. At this station, aircraft are required to remain for 11 hours at the station. Therefore, a hypothetical arc (known as a time-window arc) is created, which has a length of 11 hours. For example, the created window arc is extended from 9:00 PM to 8:00 AM the next morning. Similarly, another time-window arc may be created

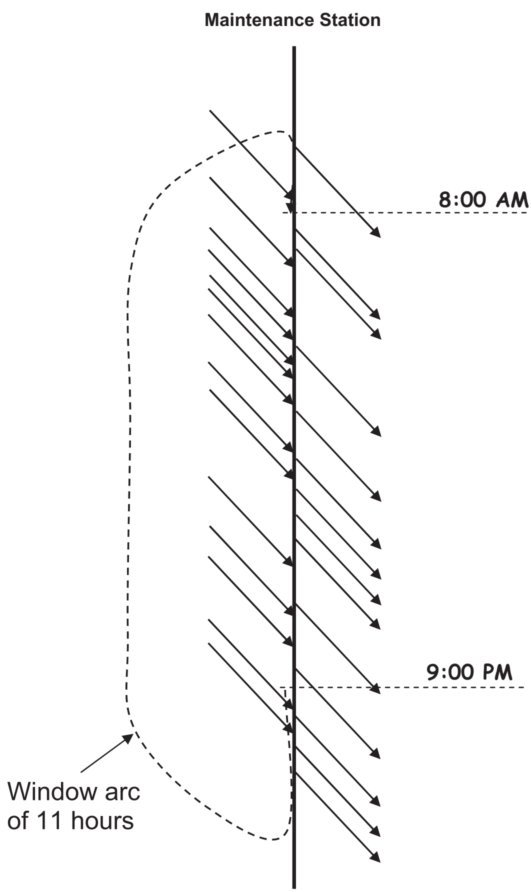


Figure 4.10 Example of a time-window arc

to extend from 11:00 PM to 10:00 AM the next morning. Then, when aircraft are forced to be on those time-window arcs, they are staying for 11 hours at the maintenance station. For each time-window requirement, a set of time-window arcs is created and added to the network. Each time-window arc should connect an interconnection node that exists before the start of the time window and another interconnection node that exists after the end of the time window, as shown in Figure 4.11. The time-window constraints can be represented as follows:

$$\sum_{w \in W_{[e,s]}} x_{we} \geq K_e^s \quad \forall e \in E, \forall s \in S \tag{4.10}$$

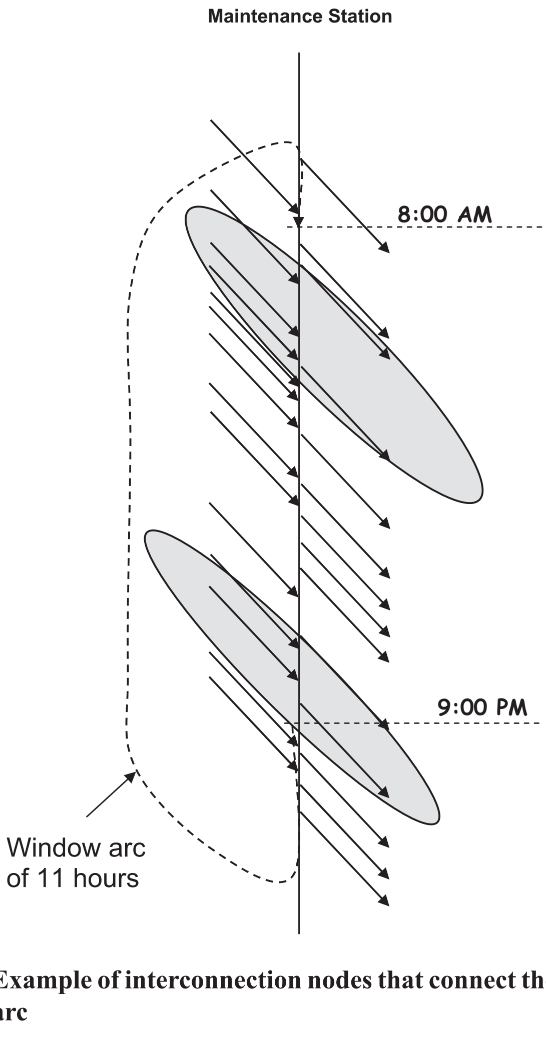


Figure 4.11 Example of interconnection nodes that connect the time-window arc

Where:

w = a time-window arc at station s

$W_{[e,s]}$ = the set of time-window arcs for aircraft type e at maintenance station s that serve as maintenance stations for aircraft type e

K_e^s = the minimum number of aircraft of type e that is required to remain at station s for a pre-specified time period.

These constraints ensure that at least the minimum number of aircraft of type e, K_e^s, stays overnight at station s for a predefined time period that is equal to the length of the time-window arc w.

Crew Constraints

The fleet assignment solution should consider a few constraints related to crew. The crew service depends on their home base (domicile) and the type of fleet that they are qualified to fly. For any air carrier, crew members are typically positioned at the air carrier's main hubs (major stations), and each crew member is allowed to fly only one type of aircraft. Several constraints have to be considered to make sure that the locations and qualifications of the crew are considered within the fleet assignment problem.

First, the solution to the fleet assignment problem should guarantee a minimum number of flying hours for the crew, which is usually specified in the crew contract. This condition can be satisfied by making sure that the total flying time by a certain aircraft type is within a pre-specified range. This range should be comparable to the available number of crew members that are qualified to fly this aircraft type. This condition can be represented mathematically as follows:

$$L_c^e \le \sum_{f \in F} \sum_{e \in E[c]} b_{fe} x_{fe} \le U_c^e \qquad \forall e \in E \tag{4.11}$$

Where:

b_{fe} = block time (or flight time) of flight f when operated by aircraft type e

L_c^e = lower bound on flying hours of crew qualified for aircraft type e

U_c^e = upper bound on flying hours of crew qualified for aircraft type e.

Second, because crew are located at different domiciles (bases), the fleet assignment solution should guarantee that, for each base, there are enough originating flights that can be assigned to the crews with different qualifications. Therefore, the model should allow a minimum number of daily departures in corresponding equipment types at each crew base. The condition can be represented mathematically as follows:

$$\sum_{f \in O[a]} \sum_{e \in E[c]} x_{fe} \ge d_e^a \qquad \forall e \in E, \forall a \in A \tag{4.12}$$

Where:

a	= crew domicile
A	= set of crew domiciles
O[a]	= set of outbound flights from domicile a
d_e^a	= minimum departures required to be assigned to aircraft type e at station a.

Third, a crew arriving on an evening flight should have a corresponding flight the next morning that departs soon after the crew's minimum required rest period. This requirement avoids long periods of crew non-utilization, which is costly to the air carrier. Figure 4.12 shows an example of fleet assignments for four flights at a small airport that has two arrivals and two departures for the air carrier. Two possible solutions to the fleet assignment problem at this station are given. In case *a*, flights F1 and F4 are assigned to aircraft type 1, and flights F2 and F3 are assigned to aircraft type 2. In case *b*, flights F1 and F2 are assigned to aircraft type 1 and flights F3 and F4 are assigned to aircraft type 2. Because there are two types of aircraft, two different crews are required. The first crew should be qualified to operate aircraft type 1, and the second crew should be qualified to operate aircraft type 2. The fleet assignment solution given in case *b* is very efficient with respect to crew utilization. The fleet assignment given in case *a* requires the crew to stay idle for long periods between flights, which is inefficient for the air carrier.

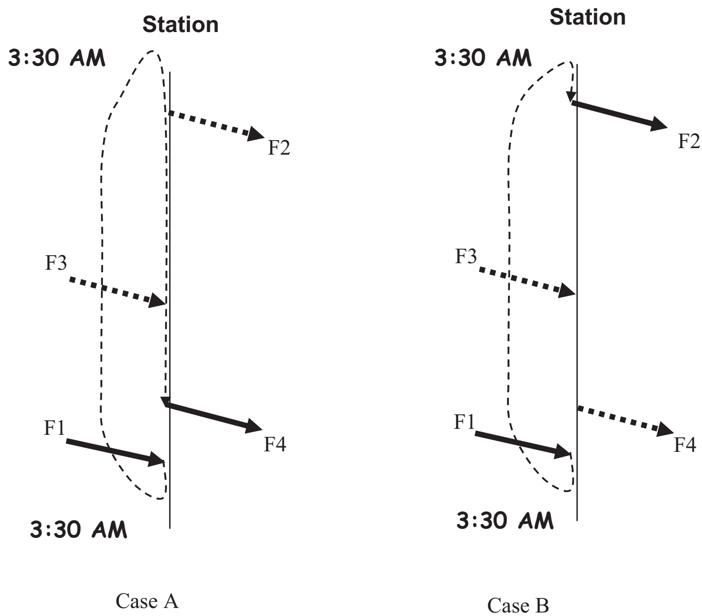


Figure 4.12 Example of fleet assignment with crew consideration

To implement this constraint, the set of terminators at station s is defined, denoted as $T[s]$. A terminator is a flight that arrives at night. Then, for each flight $t \in T[s]$, the flight set $J[t]$ is defined, which represents the set of flights originating from station s the next morning that meet the required rest interval for crew arriving by flight $t \in T[s]$. Figure 4.13 shows an example of a terminator flight $t1$ and its corresponding originating flights $J[t1]$. It should be ensured that the terminator flight is not assigned to an aircraft type unless a suitable flight has been assigned to the same aircraft type in the set of originating flights on the next day. If there is only a single flight in the terminator set, this constraint can be represented mathematically as follows:

$$\sum_{f \in J[t]} x_{fe} \geq x_{te} \quad \forall e \in E, \forall s \in S \tag{4.13}$$

This set of constraints ensures that flight t that is terminating at station s is not assigned to aircraft type e unless a suitable flight that departs after the required rest period for crew on flight t is assigned to aircraft type e the following morning. When there are multiple terminators at the station, the approach is extended by nesting the constraints of each subsequent terminator with the ones arriving before it. This nesting guarantees that the next morning a single departure will not satisfy the constraint for more than one terminator at the same time.

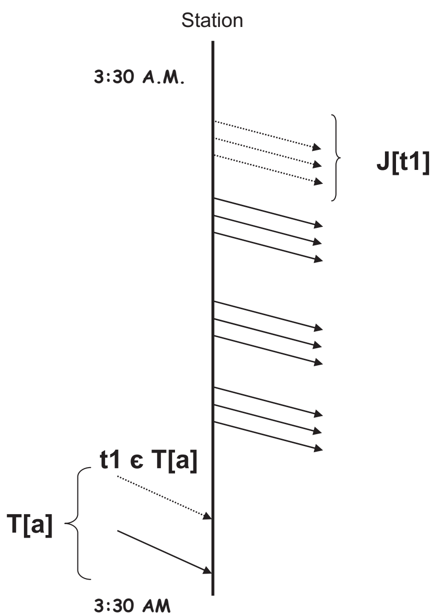


Figure 4.13 Example of a terminator flight

Other Considerations

Some other considerations might also be taken into account when solving the fleet assignment problem. These include avoiding a noisy fleet at certain airports, taking account of curfews on some fleet at some airports, the suitability of runway and gates at the airport to the aircraft type, cargo traffic, and capacity.

Primary Contributions

One of the first mathematical models to study fleet assignment models was the one developed by Abara (1989), which formulated the main constraints of the problem. However, Abara's model has several modeling and computational limitations. Subramanian et al. (1994) presents a solution procedure for the fleet assignment methodology developed by Delta Airlines. This procedure, which is referred to as Coldstart, minimizes the combination of operation and passenger spill costs subject to operational constraints. The solution strategy uses an interior point method to solve the problem initially as a linear program. Next, the structure of the original problem is modified by fixing a few variables, and the procedure solves the new problem as a mixed integer program. Hane et al. (1995) outline a model for the fleet assignment problem and discuss some of the solution issues such as degeneracy which often exists and leads to poor performance of standard linear programming (LP) solution techniques. The solution methodology employs an interior point algorithm, cost perturbation, model aggregation, branching on set-partitioning constraints, and prioritizing the order of branching. A comparison is also given on the performance of the solution procedure to standard LP-based branch and bound methodology. More details of this approach, including a complexity analysis can be found in Gu et al. (1994). Clarke et al. (1996) discuss maintenance and crew considerations in the basic daily fleet assignment problem. They also discuss several implementation issues related to the solvability of the problem. The solution methodology presented involves the use of the dual steepest-edge simplex method, in combination with a customized branch and bound strategy. Ahuja et al. (2007) present a combined fleet and through-flight assignment model, as well as a multi-criteria extension of the model to consider the planning aspects of crew and ground staff (Ahuja et al. 2003).

Smith and Johnson (2006) present an airline fleet assignment model with an imposing station purity using station decomposition. The model tries to limit the number of fleet that can operate at a single station. They demonstrate that imposing station purity on the fleet assignment model (FAM) can limit aircraft dispersion in the network and make solutions more robust relative to crew planning, maintenance planning, and operations. It is estimated that the annual net benefit of station purity is greater than $100 million for a major US domestic airline because of reduced maintenance and crew scheduling costs.

The fleet assignment models presented above are limited in that they only consider passenger demand on single flights. In reality, to get from their origin to their destination, passengers often fly on itineraries that consist of several consecutive flights by connecting at hub airports. Ignoring the demand at the itinerary-bases might lead to spilling important connecting traffic in the network. However, the fleet assignment problem becomes considerably more difficult to solve when considering itinerary-based demand. The pioneering efforts in itinerary-based fleet assignments are due to Farkas (1995). Another alternative formulation to itinerary-based fleeting is presented in Jacobs et al. (1999, 2008). The main drawbacks of the conventional fleet assignment models are given by Kniker and Barnhart (1998), who illustrate the possible impact of ignoring demand in its itinerary-based form. Barnhart et al. (2002) propose an itinerary-based model that combines a traditional fleet assignment model with a passenger mix model. Barnhart et al. (2009) offer a formulation for airline fleet assignment with enhanced revenue modeling. The formulation is motivated by the need to better model the revenue side of the objective function and presents a solution method that balances revenue approximation and model tractability. Results suggest that the approach is computationally tractable for problems of practical size, and it yields profit improvements over comparable models.

Another drawback of the conventional fleet assignment models is the assumption that the airline schedule (timetable) is known and given as an input to the fleet assignment problem. As mentioned earlier, there is a dependent relationship between setting the timetable and the fleeting. For example, an airline might decide to fly two flights between two cities with two small aircraft or, as another option, fly only one flight with a larger aircraft. Solving for the timetable and the fleet assignment mutually is a difficult problem, especially for larger-scale airlines. Rexing et al. (2000) present a fleet assignment model with time windows—small time intervals in which the flight departure times can vary slightly around a preferred time. They also demonstrate the significant benefits of their model compared to standard fixed-time fleet assignment. Ioachim et al. (1999) also propose an approach to fleet assignment and routing with time windows. Their model includes synchronization constraints to synchronize departure times of the same flight on different days of the week. Belanger et al. (2006) present a fleet assignment model with time windows, spacing constraints, and time-dependent revenues. Finally, Lohatepanont and Barnhart (2004) extend their itinerary-based fleet assignment models presented in Barnhart et al. (2002), to develop a combined scheduling and fleet assignment model. They wrap this model within another module that controls changes to the timetable.

The concept of re-fleeting or 'demand-driven dispatch' (D^3) was first introduced by Berge and Hopperstad (1993). The (D^3) concept is to resolve the fleet assignment problem with increasingly more accurate demand information as the day of operation approaches. A heuristic swap-based approach to re-fleeting is presented by Talluri (1996). Jarrah et al. (2000) demonstrate the use of the re-fleeting concept with the conventional fleet assignment models.

Another issue considered in the fleet assignment literature is incorporating consideration of operational robustness. Rosenberger et al. (2004) define short cycles as those that contain few flights, where each cycle starts at a hub. Short cycles can be canceled when disruptions occur, with minor impact on the entire schedule. Smith and Johnson (2006) present fleet assignment solutions that increase planning flexibility and reduce cost by imposing station purity, thus limiting the number of fleet types allowed to serve each airport in the schedule.

References

Abara, J. 1989. Applying Integer Linear Programming to the Fleet Assignment Problem. *Interfaces*, 19, 20-38.

Ahuja, R.K., Liu, J., Goodstein, J., Mukherjee, A., Orlin, J., and Sharma, D. 2003. Solving Multi-Criteria Combined Through-Fleet Assignment Models. In Tito A. Ciriani, Giorgio Fasano, Stefano Gliozzi, and Roberto Tadei (eds), *Operations Research in Space and Air*. Kluwer Academic Publishers, Boston, MA. pp. 233-256.

Ahuja, R.K., Liu, J., Goodstein, J., Mukherjee, A., Orlin, J., and Sharma, D. 2007. A Very Large-Scale Neighborhood Search Algorithm for the Combined Through and Fleet Assignment Model. *INFORMS Journal on Computing*, 19(3), 416-428.

Barnhart, C., Farahat, A., and Lohatepanont, M. 2009. Airline Fleet Assignment with Enhanced Revenue Modeling. *Operations Research*, 57(1): 231-244.

Barnhart, C., Kniker, T.S., and Lohatepanont, M. 2002. Itinerary-Based Airline Fleet Assignment. *Transportation Science*, 36(2), 199-217.

Belanger, N., Desaulniers, G., Soumis, F., and Desrosiers, J. 2006. Periodic Airline Fleet Assignment with Time Windows, Spacing Constraints, and Time Dependent Revenues. *European Journal of Operational Research*, 175(3), 1754-1766.

Berge, M.E. and Hopperstad, C.A. 1993. Demand Driven Dispatch: A Method for Dynamic Aircraft Capacity Assignment, Models and Algorithms. *Operations Research*, 41(1), 153-168.

Clarke, L., Hane, C., Johnson, E., and Nemhauser, G. 1996. Maintenance and Crew Considerations in Fleet Assignment. *Transportation Science*, 30, 249-260.

Farkas, A. 1995. *The Influence of Network Effects and Yield Management on Airline Fleet Assignment Decisions*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA.

Gu, Z., Johnson, E., Nemhauser, G., and Wang, Y. 1994. Some Properties of the Fleet Assignment Problem. *Operations Research Letters*, 15(2), 59-71.

Hane, C., Barnhart, C., Johnson, E., Marsten, R., Nemhauser, G., and Sigismondi, G. 1995. The Fleet Assignment Problem: Solving a Large-Scale Integer Program. *Mathematical Programming*, 70, 211-232.

Ioachim, I., Desrosiers, J., Soumis, F., and Belanger N. 1999. Fleet Assignment and Routing with Schedule Synchronization Constraints. *European Journal of Operational Research*, 119, 75-90.

Jacobs, T., Johnson, E., and Smith, B. 1999. O&D FAM: Incorporating Passenger Flows into the Fleeting Process. In R. Darrow, (ed.), *Thirty-Ninth Annual AGIFORS Symposium*, New Orleans.

Jacobs, T., Johnson, E., and Smith, B. 2008. Incorporating Network Flow Effects into the Airline Fleet Assignment Process. *Transportation Science*, 42(4), 514-529.

Jarrah, A.I., Goodstein, J., and Narasimhan, R. 2000. An Efficient Airline Re-Fleeting Model for the Incremental Modification of Planned Fleet Assignments. *Transportation Science*, 34(4), 349-363.

Kniker, T.S. and Barnhart, C. 1998. Shortcomings of the Conventional Airline Fleet Assignment Model. In *Proceedings: Tristan III*, 17-23, Puerto Rico, June 1998.

Lohatepanont, M. and Barnhart, C. 2004. Airline Schedule Planning: Integrated Models and Algorithms for Schedule Design and Fleet Assignment. *Transportation Science*, 38(1), 19-32.

Rexing, B., Barnhart, C., Kniker, T.S., Jarrah, A.I., and Krishnamurty, N. 2000. Airline Fleet Assignment with Time Windows. *Transportation Science*, 34(1), 1-20.

Rosenberger, J.M., Johnson, E.L., and Nemhauser, G.L. 2004. A Robust Fleet-Assignment Model with Hub Isolation and Short Cycles. *Transportation Science*, 38(3), 357-368.

Rushmeier, R., Hoffman, K., and Padberg, M. 1995. *Recent Advances in Exact Optimization of Airlines Scheduling Problems*. Technical Report, Department of Operations Research and Operations Engineering, George Mason University.

Smith, B. and Johnson, E. 2006. Robust Airline Fleet Assignment: Imposing Station Purity Using Station Decomposition. *Transportation Science*, 40(4), 497-517.

Subramanian, R., Scheff, R.P., Quillinan, J.D., Wiper, D.S., and Marsten, R.E. 1994. Coldstart: Fleet Assignment at Delta Airlines. *Interfaces*, 24(1), 104-120.

Talluri, K.T. 1996. Swapping Applications in a Daily Fleet Assignment. *Transportation Science*, 30, 237-248.

Chapter 5
Aircraft Routing

Introduction

The fleet assignment solution presented in the previous chapter assigns an aircraft type (for example, Boeing 737-200, Boeing 757) to each flight in the schedule. However, at this stage, it is not known yet which aircraft (defined by its tail number or nose number) is assigned to each flight. Another process is needed to assign aircraft to each flight. This process, which is known as aircraft routing, has the objective of determining the routes to be flown by each aircraft in each fleet. Therefore, each aircraft has a line of flying (LOF) that covers a set of scheduled flights. Because the fleeting solution decomposes the flight schedule into fleets, each fleet has an independent aircraft routing problem to be solved, which only considers the flights assigned to this fleet.

The LOF for each aircraft usually extends over a few days (four to seven days). It is composed of flight blocks, ground connection (aircraft turn), and maintenance activities, as shown in the hypothetical example of aircraft route presented in Figure 5.1. There are several considerations in each aircraft route that are related to aircraft turn time, maintenance, rotations, and through traffic. These considerations are discussed in more detail in the following subsections.

Aircraft Turn Time

The aircraft turn time is the time required to turn the aircraft between its arrival and its subsequent departure at any airport. The minimum aircraft turn time is required for deplaning passengers of the inbound flight, unloading cargo and baggage, aircraft cleaning, fueling, boarding passengers of the outbound flight, loading cargo and baggage, catering, and so on. The duration of the turn time depends on the aircraft type; smaller aircraft typically require a shorter turn time than larger ones. Scheduling short aircraft turn times is risky because any delay for the inbound flight typically propagates to the outbound flight. On the other hand, a lengthy aircraft turn time increases the total idle time of the aircraft, which reduces its productivity. Aircraft are air carriers' most expensive resource; accordingly, they must be utilized efficiently. Figure 5.2 shows an example of different possible turns for an inbound aircraft on flight F1. The best scenario in this example is to make the aircraft turn with flight F3. The aircraft cannot turn with flight F2 because the turn time between flight F1 and F2 is less than the minimum required turn time for this type of aircraft. The turn time between flights F1 and F4 involves a long duration of ground time for the aircraft, which is practically inefficient.

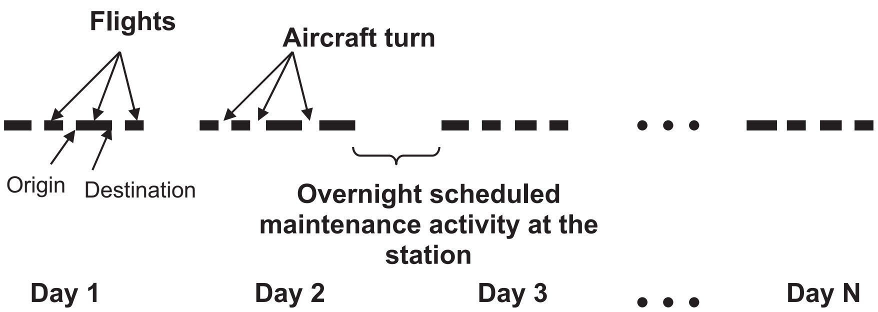


Figure 5.1 Example of an aircraft route

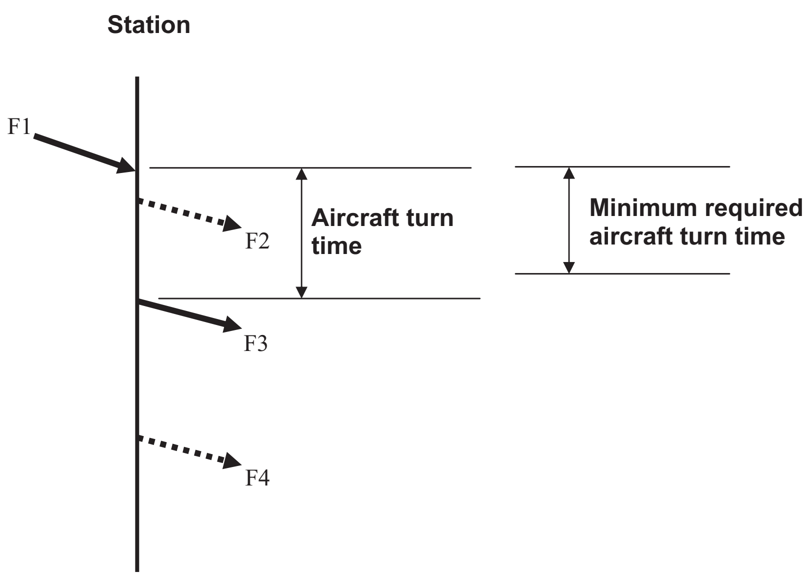


Figure 5.2 Example of a practical and efficient aircraft turn

Maintenance

In the US, the Federal Aviation Administration (FAA) requires four types of maintenance activities to be performed by commercial air carriers for their aircraft. These maintenance activities can be defined as follows:

- A-checks or line maintenance represent routine checks involving the visual inspection of the major systems of the aircraft. A-checks must be performed approximately every 65 block hours and after a certain number of takeoffs. The duration of A-checks is from three to ten hours, and they are usually performed overnight in the US domestic market.

- B-checks involve a detailed visual inspection of most parts of the aircraft and are performed once every several months.
- C- and D-checks require that the aircraft be taken out of service to a hangar for about a month to be thoroughly inspected. Most of the perishable systems are upgraded. These checks are performed once every one to four years.

Aircraft routing is typically more concerned with the A-checks, where a typical aircraft route might require one or two breaks for maintenance activities. Maintenance activities can only be performed at specific stations, known as maintenance stations. Maintenance stations are supported by resources, equipment, and personnel that have predefined operations capacity.

It should be mentioned that the fleet assignment solution considered in the previous chapter includes several considerations for aircraft maintenance. These considerations specify that a minimum number of overnight aircraft of a particular type are at the different maintenance stations. Also, the number of overnight aircraft that remain at the maintenance station should be comparable to the maintenance capacity of the station. Furthermore, the minimum time duration that aircraft should remain at the maintenance station should be considered. However, the fleet assignment problem does not influence the spacing of the maintenance visits for each tail number. Adding constraints to take into account maintenance visit spacing would make the fleet assignment problem computationally intractable. Therefore, spacing maintenance visits evenly for each tail number is one of the main considerations of the aircraft routing problem.

Rotation

To explain the main considerations in aircraft rotation, consider the hypothetical flight schedule in the domestic US market, given in Figure 5.3 (Clarke et al. 1997). Flights 1 and 3 are from Dallas to Chicago and flights 2 and 4 are from Chicago to Dallas. It is assumed that this schedule is repeated on a daily basis at the same departure times, and all aircraft start from Dallas. Several possibilities for aircraft rotations may be considered as follows.

Rotation 1 An aircraft A operates flights 1 and 2 on the first day and flights 3 and 4 on the second day. This aircraft makes an overnight stop in Dallas. This overnight stay requires the existence of another aircraft, B, which operates flights 3 and 4 on the first day and flights 1 and 2 on the second day. Aircraft B also makes the overnight stay in Dallas. Both aircraft A and B are flying the rotation 1–2 ◦ 3–4 ◦ 1, where the symbol ◦ represents an overnight stay between flights. This rotation covers the four flights by two aircraft. The main characteristic of this rotation is that the aircraft used to satisfy this schedule must cover each flight in the schedule. Both aircraft A and B must also cover all flights in the same order.

Rotation 2 The rotation 1–4 ◦ 3 ◦ 2 ◦ 1 requires that three aircraft cover the same four flights. Positioning flight 4 after flight 1 requires a long ground time in Chicago. Hence, two other aircraft are needed to fly the flights 2 and 3 that are scheduled at the same time from two different stations.

Rotation 3 An aircraft A operates flights 1 and 2 everyday and makes an overnight in Dallas. This rotation requires the existence of another aircraft B, which operates flights 3 and 4 everyday. Aircraft B also makes the overnight in Dallas. In this situation, flights 1 and 2 are always covered by aircraft A, while flights 3 and 4 are always covered by aircraft B. When flights covered by one aircraft are separate from flights covered by another aircraft in the fleet, it is called a broken rotation. The coverage of rotation 3 can be represented as 1–2 ◦ 1 ◊ 3–4 ◦ 3, where the symbol ◊ indicates where coverage breaks into different sets of aircraft. This break is sometimes called a continuity break because it breaks the continuous flow of an aircraft around a cycle that contains all the flight legs. There can be several continuity breaks in a rotation and, within each break, there can be many overnights. It is common practice on the part of air carriers to avoid broken rotations (or continuity breaks).

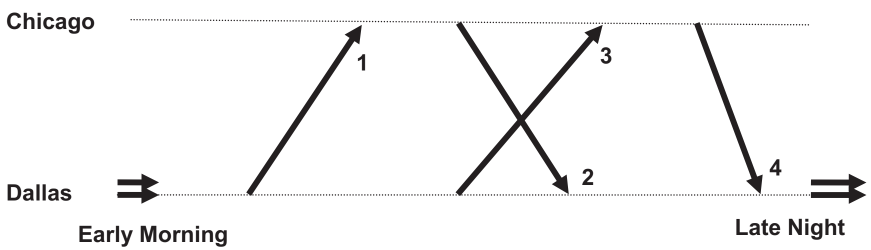


Figure 5.3 Example of aircraft rotation

Through Traffic

As defined earlier in the previous chapter, through traffic is the traffic traveling from point A to point B through point C without an aircraft change at point C. To explain the relationship between through traffic and aircraft routing, consider the two flight arrivals and two flight departures at the Atlanta airport shown in Figure 5.4. Assume also that there are two aircraft to operate these four flights.

Two possible flight rotations exist for these two aircraft. In the first rotation, the first aircraft operates on the route Seattle-Atlanta-Orlando, and the second aircraft operates on the route Boston-Atlanta-Miami. In this case, we assume that there are x passengers connecting from Seattle to Orlando, and y passengers connecting from Boston to Miami. In the second rotation, the first aircraft operates on the route Seattle-Atlanta-Miami, and the second aircraft operates on the route Boston-Atlanta-Orlando. We assume that there are m passengers connecting from Seattle to Miami, and n passengers connecting from Boston to Orlando. If through

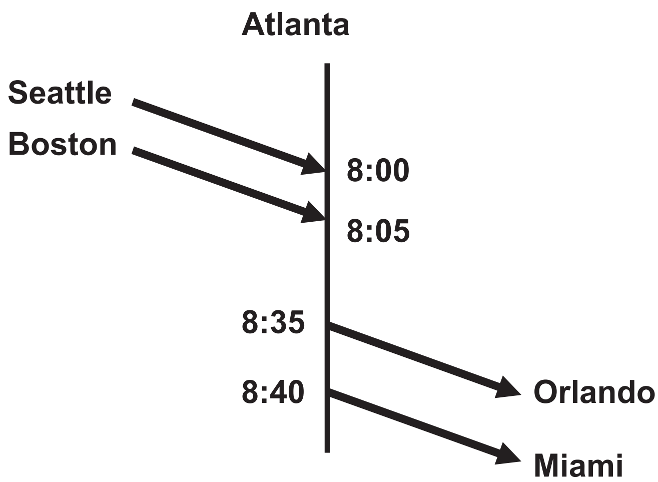


Figure 5.4 Example of through traffic

traffic is to be considered in aircraft routing, the first rotation is considered if $x + y > m + n$; otherwise, the second rotation may be considered.

Solution Methodology

As mentioned earlier, the main objective of the aircraft routing problem is to determine the routes to be flown by each aircraft in the fleet. One of the earliest formulations of the problem is the set partitioning (Kabbani and Patty 1992) which is commonly used in scheduling problems where there are several scheduled tasks (flights) and several resources (aircraft), and every task is to be covered by only one resource. To explain the set partitioning formulation, consider the set of matrices presented in Figure 5.5. The first matrix on the left, which has a size [I × 1] includes the list of flights to be covered in aircraft rotations. Each flight $i \in I$ is to be covered in the rotation of one and only one aircraft. The right-hand-side matrix has a size of [I × 1]. Each element in this matrix, b_i associated with flight i, is equal to one, which indicates that flight i has to be included in one and only one aircraft route. The matrix in the middle, which is known as matrix A, is of size [I × J] and represents possible aircraft routes. Each route $j \in J$, which is represented by a column, consists of a set of flights that can connect in a feasible aircraft route. An item in this matrix, a_{ij}, is equal to one, if flight i is part of aircraft route j and zero, otherwise. Each route in this matrix has a cost function c_j that measures the quality of the aircraft route. The solution to the set partitioning problem is to find the most efficient aircraft routes that cover all flights in the schedule. Accordingly, the decision variable x_j is defined, which is equal to one if route j is selected in the solution and zero,

Line of work (aircraft route)

a_{ij}

$$
\begin{bmatrix} Flight \\ i=1 \\ i=2 \\ i=3 \\ \\ \\ \\ \\ I \end{bmatrix}
\quad
\begin{bmatrix}
j=1 & j=2 & j=3 & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & j=J \\
1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\
0 & 1 & 0 & & & & & & & & & & & & 0 \\
1 & 0 & 0 & & & & & & & & & & & & 0 \\
1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 \\
1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\
0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\
1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0
\end{bmatrix}
\quad
\begin{bmatrix} b_i \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}
$$

Figure 5.5 Representation of the aircraft routing solution

otherwise. The objective is to find a LOF for each aircraft (or tail number) $t \in T$ where T is the set of available aircraft.

Problem Formulation

The set partitioning problem can be represented mathematically as follows:

$$Minimize \quad \sum_{j \in J} c_j x_j \tag{5.1}$$

Subject to:

$$\sum_{j \in J} a_{ij} x_j = b_i \qquad \forall i \in I \tag{5.2}$$

$$\sum_{j \in J} x_j = 1 \qquad \forall t \in T \tag{5.3}$$

$$x_j \in \{0,1\}, \qquad j = 1, 2, \ldots, J \tag{5.4}$$

Where:

$$a_{ij} = \begin{cases} 1 & \textit{if flight leg "i" is part of aircraft route "j"} \\ 0 & \textit{otherwise} \end{cases}$$

The objective function is to minimize the total cost associated with the selected aircraft route. The first set of constraints is to ensure that each flight is covered by one and only one aircraft routing. The second set of constraints is to make sure that each aircraft is assigned to only one feasible routing. The third set of constraints is to ensure the integrality of the decision variables.

If there is a limited number of aircraft (that is, too many flights in the schedule), the given formulation may produce infeasibility. To overcome this infeasibility, dummy aircraft routes (columns) are added to the end of matrix A. Each dummy column represents an aircraft route that consists of a single flight. The new form of matrix A is as shown in Figure 5.6. A very high cost c_j is to be associated with each dummy column, so a dummy aircraft is included in the solution only when there is no aircraft to cover the flight. Accordingly, when the dummy column is selected in the set partitioning solution, this indicates that the corresponding flight cannot be covered in any of the routings.

Solution Algorithm

The solution method consists of two main steps, which are known as column generation and optimization. The main objective of column generation is to generate a reasonably large set of feasible aircraft routes (columns). The column generation procedure is usually performed using an exhaustive enumeration. Flights are sequenced together in a single column according to a set of pre-specified rules. These rules are based on the minimum required turn time for the aircraft, through traffic, number of block hours or number of flights before a maintenance activity, avoiding continuity breaks, and so on. For air carriers that operate in a hub-and-spoke network structure where there are many possibilities of flight connections, the enumeration usually results in a large number of columns, as shown in Figure 5.7.

Dummy columns

$$
\begin{bmatrix} Flight \\ i=1 \\ i=2 \\ i=3 \\ \\ \\ \\ \\ I \end{bmatrix}
\quad
\left[\begin{array}{ccccccccccccccc|cccccccc}
j=1 & j=2 & j=3 & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & j=J & i=1 & i=2 & i=3 & .. & .. & .. & .. & i=I \\
1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & & & & & & & & & & & & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & & & & & & & & & & & & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{array}\right]
\quad
\begin{bmatrix} b_i \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}
$$

Figure 5.6 Representation of the aircraft routing solution with dummy variables

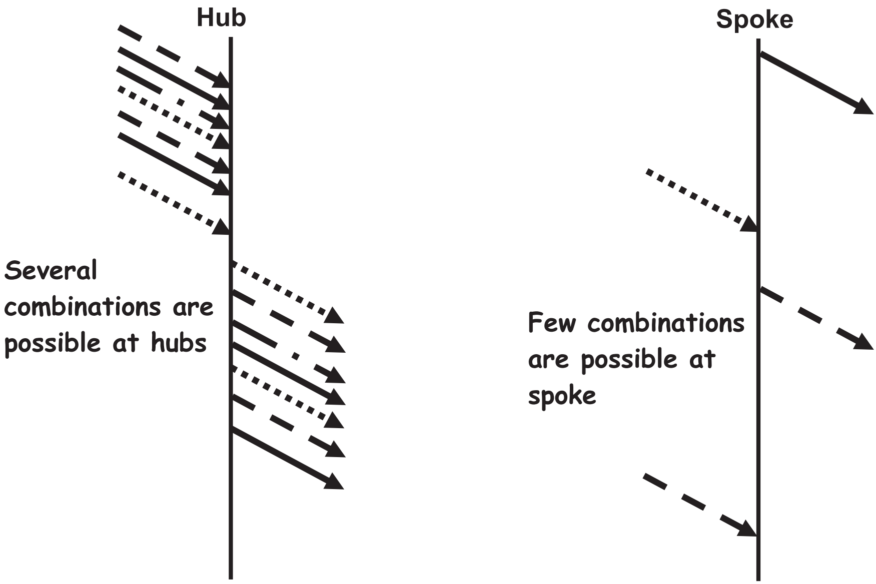


Figure 5.7 Many connection possibilities at the hub station compared to a spoke

The objective of optimization is to determine the combination of aircraft routes that cover all flights. The optimization problem is not necessarily straightforward in the case of a larger problem. The set partitioning problem is known to be a non-deterministic polynomial time (NP)-hard problem, where the problem size increases exponentially with the number of flights. When the number of flights increases, heuristics are used to solve the optimization problem. To find the solution for the problem, optimization software is used.

Primary Contributions

One of the earliest studies of aircraft routing was undertaken by Daskin and Panayotopoulos (1989), who formulate a route selection problem to maximize profits in a single hub-and-spoke network as a Mixed Integer Program (MIP). Feo and Bard (1989) study the maintenance location problem which involves estimating the minimum number of maintenance stations required to support the proposed flight schedule. They use the one-day routings between overnight cities as input and formulate the maintenance base location problem as a minimum cost, multi-commodity network flow problem with integer restrictions on the variables. Kabbani and Patty (1992) study the maintenance routing problem for American Airlines. The problem is formulated as a set partitioning model where a column represents a possible week-long routing and a row represents a

flight. Routings with characteristics such as violation of the connection times, maintenance violations, and continuity breaks are penalized using predefined costs. This formulation proves to be too large to solve for a large fleet. Accordingly, the problem is separated into two sub-problems. First, they solve for the appropriate 'over-the-day' routings. Next, they solve for the connections among these routings. Flight-swapping options are considered when the sub-problem approach does not find an appropriate solution. Clarke et al. (1997) present a mathematical formulation for the aircraft rotation problem in which a specific route for each aircraft is determined under the maintenance feasibility constraints. This formulation is similar to the asymmetric traveling-salesman problem. The aircraft routing problem is solved by Lagrangian relaxation and sub-gradient optimization, and computational results on real data from a major airline are presented. Gopalan and Talluri (1998) develop a polynomial-time algorithm for long-term aircraft maintenance routing. The algorithm considers a scheme of swaps and interchanges among aircraft to form a feasible solution to modify routings. Initially, a set of sequences of flight legs are generated. Each sequence is constructed by fixing all connections during the day at non-overnight stations using simple rules such as first-in-first-out or last-in-first-out. Accordingly, the obtained network contains only overnight stations and the fixed sequences between these stations. Then the maintenance on this reduced network is routed to try to meet a multi-day maintenance requirement for each aircraft. Next, aircraft swapping or fleet-swapping options are considered to improve maintenance routing and eliminate any locked rotations.

Sriram and Haghani (2003) present a model for maintenance scheduling and aircraft reassignment. In this model, the maintenance scheduling is performed after aircraft have been assigned to routes. The problem is formulated as an integer program. A heuristic depth-first search and random search methods are used to solve the resulting integer program. The authors show that this approach produces results, which compare well with optimal results obtained with branch-and-bound.

Sarac et al. (2006) consider the aircraft routing problem on an operational level rather than on a planning level. They develop a formulation for an operational aircraft maintenance routing problem that includes maintenance resource availability constraints. A branch-and-price algorithm for solving this problem is proposed. In this algorithm, a modification is made for the branch-on, follow-on branching rule typically used for solving similar problems, due to the resource constraints. The efficiency of this solution approach is tested under a combination of heuristic choices for column generation, and selection is investigated. Grönkvist (2005, 2006) presents a recent complete solution approach for the aircraft routing problem. In this approach, constraint programming is combined with column generation and local search. Grönkvist claims that this approach can meet both the running time requirements of operations management and the quality requirements of long- and mid-term planning.

References

Clarke, L.W., Johnson, E.L., Nemhauser, G.L., and Zhu, Z. 1997. The Aircraft Rotation Problem. *Annals of Operations Research*, 69, 33-46.

Daskin, M.S. and Panayotopoulos, N.D. 1989. A Lagrangian Relaxation Approach to Assigning Aircraft to Routes in Hub and Spoke Networks. *Transportation Science*, 91-99.

Feo, T.A. and Bard, J.F. 1989. Flight Scheduling and Maintenance Base Planning. *Management Science*, 35, 1415-1432.

Gopalan, R. and Talluri., K.T. 1998. The Aircraft Maintenance Routing Problem. *Operations Research*, 46(2), 260-271.

Grönkvist, M. 2005. The Tail Assignment Problem. Ph.D thesis, Department of Computing Science, Chalmers University of Technology, Gothenburg, Sweden.

Grönkvist, M. 2006. Accelerating Column Generation for Aircraft Scheduling using Constraint Propagation. *Computers and Operations Research*, 33, 2918-2934.

Kabbani, N.M. and Patty, B.W. 1992. Aircraft Routing at American Airlines. In *Proceedings of the Thirty-Second Annual Symposium of AGIFORS*, Budapest, Hungary.

Sarac, A., Batta, R., and Rump, C.M. 2006. A Branch-and-Price Approach for Operational Oriented Aircraft Maintenance Routing. *European Journal of Operational Research*, 175(3), 1850-1869.

Sriram, C. and Haghani, A. 2003. An Optimization Model for Aircraft Maintenance Scheduling and Reassignment. *Transportation Research Part A*, 37, 29-48.

Chapter 6
Crew Planning

Introduction

The crew cost is the second highest operation cost for the air carrier after fuel. Efficient crew scheduling has a significant impact on the air carrier's operation cost and crew productivity. The crew workload is usually determined by: 1) regulations of the aviation authority—that is, the FAA in the US—which are set to guarantee safe operations; 2) agreements between the air carrier and the crew union, which guarantee good quality of life measures for the crew; and 3) other measures that are determined by the air carrier to improve its service quality.

The aviation authority regulations guarantee safe operation by making sure that crew members do not operate while exhausted or under strenuous conditions. In that regard, the aviation authority ensures that crew members get enough connection time between flights, enough layover periods, and a reasonable number of flying or working hours every day. The aviation authority also regulates the minimum number of crew members to be staffed on each flight. The crew members, through their union, may negotiate minor work rules that are related to the start and end time of each duty period, vacation days, and so on. Finally, air carriers can make additional enhancements to the work rules of their crew members. For example, an air carrier may increase the number of flight attendants on each flight to offer better service to passengers or provide longer layover or shorter duty periods for its crew to improve their working standards.

Cockpit crew members are typically trained and certified to fly one type of aircraft fleet. Due to safety requirements, they cannot fly any other fleet at the same time. In addition cockpit crew members are ranked based on their seniority from captain, to first officer (F/O), and in some cases second officer (S/O). Most aircraft fleets are operated by a captain and a F/O. Thus, each cockpit crew member is defined by their fleet type and position. Similarly, flight attendants are trained to serve on either wide-body or narrow-body aircraft. Most air carriers allow flight attendants to operate one of these two aircraft types.

The crew planning process is divided into two sequential phases known as crew pairing and crew rostering (or assignment). First, pairings (or crew rotations) are formed on an anonymous basis. The objective of crew pairing is to satisfy the crew needs on each flight by forming a sequence of flights known as pairings or trippairs out of the flight legs. A trippair usually extends over two to seven days, and the crew member has to return to their base at the end of it. Second, in the crew rostering, the pairings are sequenced with possible other activities such as vacation, training duties, and ground duties to form monthly rosters to be assigned to the individual crew members.

For most major domestic air carriers in the US, trippairs are constructed on a monthly basis. Then, a monthly LOF is constructed out of trippairs for each crew member. The LOF considers the crew vacation and training requirements. The crew members are then assigned to these monthly lines either by auctions or based on seniority. On a monthly basis, the crew members of an air carrier can be classified into line holders and reserve. Line holders, as the name implies, are crew members who are assigned a LOF during the month. Reserve crew members are used when needed to cover open trippairs and open flights. Some of the reserve crew members are assigned standby duties at airports, where they stay at the airport for a predefined period (about four hours). During this period, the standby crew should be ready to operate any flight that has an open position, which usually results from schedule irregularities.

Both the crew pairing and rostering processes involve satisfying many complex working rules, regulations, and crew preferences, while minimizing the overall cost of the crew. While cost is an important aspect, determining the crew rostering also considers the quality of life of the crew as related to the distribution of work, training, and vacation activities over the month.

Crew Pairing

Trippair Definition

A trippair is a sequence of flights (segments), the first of which starts at the crew domicile (home base) and the last of which ends at the crew domicile. Each segment consists of a single flight takeoff and landing with no intervening stops. A trippair typically extends over two to seven days, and consists of successive sets of duty periods and rest (layover) periods. Figure 6.1 presents an example of a typical trippair. As shown in the figure, flights are grouped into duty periods which each, typically, extends over one day. The crew's duty period starts one hour before the

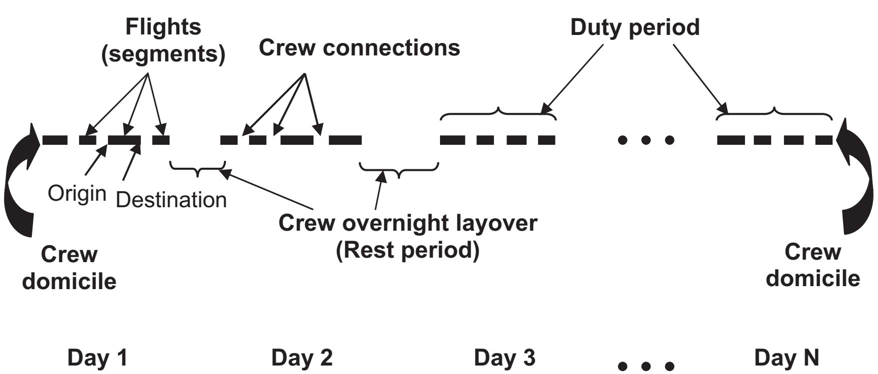


Figure 6.1 Example of a typical crew trippair

departure time of the first flight in the duty period (briefing period) and ends about 15 minutes after the arrival time of the last flight. The crew's minimum connection period, maximum duty period, and the minimum layover period are determined on the basis of regulations set by the aviation administration to guarantee safe operation and good quality of life for the crew.

Crew Work Rules

The crew's restrictive work rules make crew pairing extremely complex. Examples of the principal crew working rules can be summarized as follows:

- The maximum time on duty should not exceed a predefined maximum threshold (about 13 hours).
- The total flying time within a duty period should not exceed a predefined maximum threshold (about eight hours).
- Crew members may be paid an overtime rate when they fly more than five hours consecutively.
- At any time, the maximum flying time in the previous 24 hours should not exceed a predefined threshold (usually eight hours).
- There is a predefined threshold on the minimum rest (layover) periods that are given between two duty periods (usually ten hours).
- When flying time in the last 24 hours has exceeded more than eight hours, the next rest period should not be less than 12 hours.
- Layover time cannot exceed a maximum amount (usually 36 hours).
- The ground time between two flights in the same duty period (turn time) should be greater than a predefined minimum, which typically depends on the airport size at which the connection is made and whether it is a domestic or international airport.
- Crew turn time cannot exceed a predefined maximum amount (typically four hours).
- The pay of the crew member is determined on the basis of duty time, flying time, and time away from base.
- The crew has a minimum guaranteed pay-per-duty period.
- The crew has a guaranteed percentage of duty time that counts as flying time.
- The crew is paid per diem for each layover period away from their home base.
- The crew is paid bonuses if their total time away from base exceeds a predefined maximum period.
- The crew is paid bonuses if their trippair begins at one co-terminal and ends at another. Co-terminals are airport stations that are close to each other (that is, within a driving distance) such as John Kennedy (JFK), Newark (NWK), and LaGuardia (LGA) airports in New York.

The above crew legalities make the optimization of crew trippairs very complex. In addition, the hub-and-spoke network structure, which is adopted by most major air carriers, further complicates the problem due to the increased number of possible connections between flights at hubs, as shown in Figure 6.2. In the next section, the mathematical formulation of the crew trippair optimization problem is presented.

Crew Pairing Problem Formulation

As with the aircraft routing problem, the crew pairing problem is formulated as a set partitioning problem. As mentioned earlier, the set partitioning formulation is used in scheduling problems when several scheduled tasks (flights) need to be covered by several available resources (crew), and every task is to be covered by only one resource. The crew trippair problem is divided into several similar problems, where one problem is to be solved for each fleet and crew position on each flight (captain, F/O, and so on).

Similar to the description given in the previous chapter, consider the set of matrices presented in Figure 6.3. The first matrix on the left, which has a size [I × 1] includes the lists of flights to be covered in crew trippairs. Each flight $i \in I$ is to be covered in the trippair of one and only one crew member. The right-hand-side matrix has a size of [I × 1]. Each element in this matrix b_i associated with flight i is equal to one, which indicates that flight i has to be included in one and only one trippair. The matrix in the middle, which is known as matrix A, is

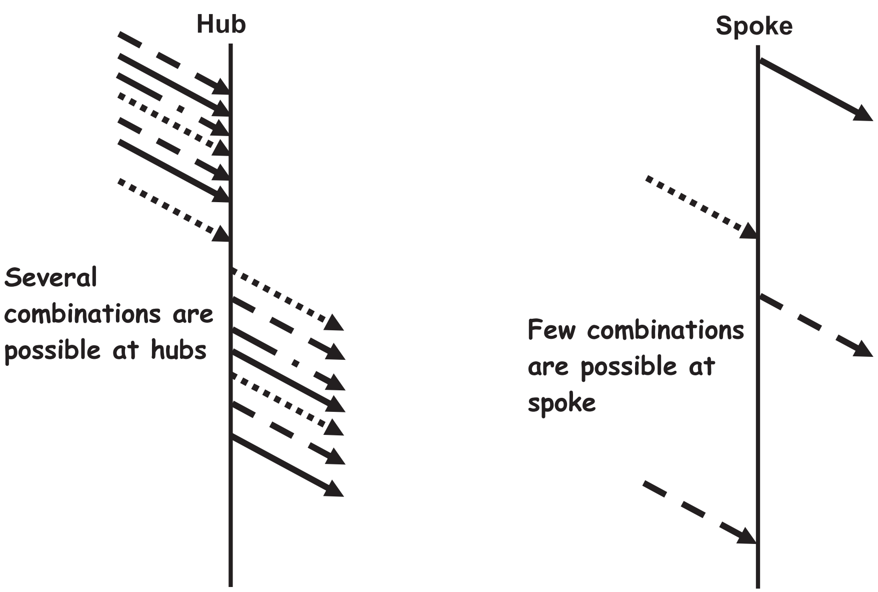


Figure 6.2 Possibilities of crew connections at hub and spoke

Crew trippair

$\mathbf{a_{ij}}$

$$
\begin{bmatrix} Flight \\ i=1 \\ i=2 \\ i=3 \\ \\ \\ \\ \\ I \end{bmatrix}
\quad
\begin{bmatrix}
j=1 & j=2 & j=3 & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & j=J \\
1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\
0 & 1 & 0 & & & & & & & & & & & & 0 \\
1 & 0 & 0 & & & & & & & & & & & & 0 \\
1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 \\
1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\
0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\
1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0
\end{bmatrix}
\quad
\begin{bmatrix} b_i \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}
$$

Figure 6.3 Representation of the crew pairing solution

of the size [I × J] and represents the set of possible crew trippairs. Each route $j \in J$, which is represented by a column, consists of a set of flights that can connect in an operationally feasible and legal crew trippair. Associated with each trippair is a cost element c_j that gives the total cost of each crew trippair. An item in this matrix a_{ij} is equal to one if flight i is part of trippair j, and zero otherwise. The solution to the set partitioning problem is to find the most efficient trippairs in matrix A that cover all flights in the schedule. Therefore, a variable x_j is defined, which is equal to one if trippair j is selected in the solution, and zero otherwise.

The set partitioning problem can be represented mathematically as follows:

$$Minimize \quad \sum_{j \in J} c_j x_j \tag{6.1}$$

Subject to:

$$\sum_{j \in J} a_{ij} x_j = b_i \qquad \forall i \in I \tag{6.2}$$

$$x_j \in \{0,1\} \qquad j = 1, 2, ..., J \tag{6.3}$$

Where,

$$a_{ij} = \begin{cases} 1 & if\ \ flight\ leg\ "i"\ is\ part\ \ of\ trippair\ \ "j" \\ 0 & \qquad\qquad otherwise \end{cases}$$

The objective function is to minimize the total cost associated with the selected crew trippairs that cover the flights. The first set of constraints is to make sure

that each open position (for example, captain or F/O) on the flight is covered by one and only one crew member. The second set of constraints is to ensure the integrality of the decision variables.

Similar to the aircraft routing formulation presented in the previous chapter, if there are too many flights in the schedule, the previous formulation may produce infeasibility. To take this into account, dummy trippairs (columns) are added to the end of matrix A. Each dummy column represents a trippair that consists of a single flight. The new form of matrix A is shown in Figure 6.4. A very high cost c_j is to be associated with each dummy column so that these dummy trippairs are not included in the solution. Accordingly, if any of the dummy columns are selected in the set partitioning solution, the corresponding flight cannot be covered in any of the trippairs.

The solution method is composed of two main steps that are known as column generation and optimization. The main objective of column generation is to generate a reasonably large set of feasible and legal crew trippairs. The column generation procedure is usually performed using exhaustive enumeration. Flights are sequenced together in a single column according to a set of pre-specified rules that satisfy the crew work rules identified above. For air carriers that operate in a hub-and-spoke network structure, enumeration usually results in a large number of columns; with many possibilities of flight connections at the hubs, as shown in Figure 6.2. Given that optimal trippairs represent a subset of all generated trippairs, it is very important to generate a reasonably large number of columns to guarantee finding a close-to-optimal solution within these columns.

The major difficulty with this formulation is that the column generation procedure produces a significant number of columns that make the optimization problem hard to solve. A test case on a fleet of 600 flights would generate an estimated million legal duty periods. Because four- to five-day trippairs are possible for this fleet, it is easy to see that the number of legal trippairs can rise to

Dummy columns

$$\left[\begin{array}{c} \mathit{Flight} \\ i=1 \\ i=2 \\ i=3 \\ \\ \\ \\ \\ I \end{array}\right] \quad \left[\begin{array}{ccccccccccccccc|cccccccc} j=1 & j=2 & j=3 & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & j=J & i=1 & i=2 & i=3 & .. & .. & .. & .. & i=I \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & & & & & & & & & & & & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & & & & & & & & & & & & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \quad \left[\begin{array}{c} b_i \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{array}\right]$$

Figure 6.4 The trippairs matrix with the set of dummy trippairs

astronomical figures. The optimization step is to determine, from the combination of trippairs selected, the solution that can cover all flights. The optimization problem is not necessarily straightforward in the case of a larger problem. The set partitioning problem is known to be a NP-hard problem, where the problem size increases exponentially with the number of flights. When the number of flights increases, heuristics are used to solve the optimization problem.

Solution Methodology for Large-scale Problems

For large-scale airlines, a large number of crew trippairs are generated to cover the available flights. With existing computer capabilities, the solution of the set partitioning formulation cannot be obtained in one step by using any of the available optimization software. A solution algorithm for this problem has been developed by Rubin (1973), and this is used by most US air carriers to drive the final crew schedule. Instead of attempting to solve a set partitioning problem with a large matrix, this algorithm involves systematically extracting thousands of very small set partitioning problems from the original large one. These small problems are solved consecutively, and in each solution a small improvement is produced for the original problem.

The main steps of Rubin's algorithm can be described as follows:

Consider that:

S = set of legal trippairs (columns) that covers all flights once

P_j, $j = 1,2,\ldots,r$ = a trippair in the solution

$F_j = F(P_j)$ $j = 1,2,\ldots,r$ = set of flights (rows) covered by trippair P_j

F = total set of flights

c_j, $j = 1,2,\ldots,r$ = cost of trippair P_j.

Step 1 It is assumed that an initial fully legal solution S (the set of legal trippairs that covers all flights once) is given. This set can be determined manually or by any automated method. One possible approach to finding this initial feasible solution is to assume a set S in which each trippair covers only one flight. This solution is an expensive initial solution, but feasible. The initial solution can be represented mathematically as follows:

$$S = \{P_j\} \quad j = 1,2,\ldots,r \tag{6.4}$$

$$F = \bigcup_j F_j \tag{6.5}$$

$$F_k \bigcap F_j = \phi, \quad j \neq k \tag{6.6}$$

The set S consists of a set of legal trippairs, as given in equation 6.4. The set of all flights F is the sum of all flights in all trippairs, as given in equation 6.5.

Finally, equation 6.6 indicates that every pairing in the set S covers a different set of flights. Figure 6.5 gives an example of the trippairs matrix, where each flight is only covered by one trippair.

Step 2 Trippairs (columns) are sorted in a descending order according to their cost, as shown in Figure 6.6. Therefore, the trippairs are located in the matrix according to the following rule:

$$c_k \geq c_j, \text{ if } k < j \tag{6.7}$$

Step 3 A subset of columns $T \subset S$ is selected, as shown in Figure 6.7. The set T consists of one or more columns at a time. The cost $C_{current}$ of the trippairs in the subset T is as follows:

$$C_{current} = \sum_{p_j \in T_{current}} c_j \tag{6.8}$$

The set of flights that are covered by the trippairs in the set T are identified as follows:

$$F = \bigcup_{P_j \in T} F_j \tag{6.9}$$

Step 4 (sub-problem) For the flights in set F_T (that is, flights that are covered by the trippairs in the set T), generate all legal trippairs that cover these flights.

Solve the set partitioning problem to find the optimal trippairs that cover the set of flights F_T. The cost of this new solution to the sub-problem can be represented as follows:

$$C_{new} = \sum_{p_j \in T_{new}} c_j \tag{6.10}$$

$$\begin{bmatrix} Flight \\ i=1 \\ i=2 \\ i=3 \\ \\ \\ \\ \\ I \end{bmatrix} \quad \begin{bmatrix} j=1 & j=2 & j=3 & .. & .. & .. & .. & j=I=J \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} b_i \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Figure 6.5 Example of the trippairs matrix, where each flight is only covered in one trippair

$$
\begin{array}{ccc}
 & \text{Highest-cost itinerary} \qquad \text{Lowest-cost itinerary} & \\
\begin{bmatrix} Flight \\ i=1 \\ i=2 \\ i=3 \\ \\ \\ \\ \\ I \end{bmatrix} &
\begin{bmatrix}
j=1 & j=2 & j=3 & .. & .. & .. & .. & j=I=J \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0
\end{bmatrix} &
\begin{bmatrix} b_i \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}
\end{array}
$$

Figure 6.6 Example of the trippairs matrix, where each flight is only covered in one trippair, and the trippairs are sorted based on their cost

$$
\begin{array}{ccc}
 & \overbrace{\qquad\qquad\qquad\qquad}^{\text{Set } T} & \\
\begin{bmatrix} Flight \\ i=1 \\ i=2 \\ i=3 \\ \\ \\ \\ \\ I \end{bmatrix} &
\begin{bmatrix}
j=1 & j=2 & j=3 & .. & .. & .. & .. & j=I=J \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0
\end{bmatrix} &
\begin{bmatrix} b_i \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}
\end{array}
$$

Figure 6.7 Selecting a subset *T* from the current trippairs matrix

Step 5 If $C_{new} < C_{current}$, then the set of trippairs obtained from solving the sub-problem (step 4) is better than the set of trippairs in the original set T ($T_{current}$). In this case, the subset $T_{current}$ is replaced by the subset T_{new} in the matrix *S*, thus obtaining a new improved solution to the full problem.

Step 6 In any case, a new subset *T* can be selected, and steps 3, 4, and 5 are repeated until a final acceptable solution is obtained.

Alternative Approach

In the previous approach to solving the crew scheduling problem, the focus is on selecting a set of columns (subset T) for each iteration (column approach). Other alternative approaches are called hybrid approaches, in which alternate iterations look at the column structure or the row structure, as shown in Figure 6.8. In every other iteration, some of the rows (flights) are selected, and all feasible trippairs that cover these flights are generated. Then, the set partitioning optimization model is solved to find the optimal set of trippairs that covers these flights. The other steps are similar to the column approach presented by Rubin (1973).

Deadhead Flights

A crew member is defined as a deadhead when they are traveling on a flight as a passenger as part of their duty. The crew is usually deadheaded from one station to another when there is an imbalance between the number of flights and the number of crew members at the different stations. The deadhead flights are used to reposition the crew to stations that need them or to return the crew back to their domicile stations. Usually, time spent on deadhead flights is considered to be a part of the crew's duty period. However, it is not considered part of their flying time.

To generate trippairs that include deadhead flights, the column generation process must be adjusted. In Rubin's approach, the deadhead flights can be included by appending a set of dummy rows to the matrix to represent the possible deadhead flights. Therefore, there are two sets of flights included: the original set of flights and the dummy set of flights. The main drawback of this approach is that users must specify the flights to be considered as deadhead flights.

When a subset of the original flights is selected at any iteration of Rubin's algorithm, the dummy flights can be used to generate columns that cover the

$$
\begin{array}{l}
\text{Set of flights to be} \\
\text{considered in the} \\
\text{hybrid approach}
\end{array}
\left\{\begin{bmatrix} \textit{Flight} \\ i=1 \\ i=2 \\ i=3 \\ \\ \\ \\ \\ I \end{bmatrix}\right.
\quad
\begin{bmatrix}
\overbrace{j=1 & j=2 & j=3}^{\text{Set } T} & .. & .. & .. & .. & j=I=J \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\quad
\begin{bmatrix} b_i \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}
$$

Figure 6.8 Trippairs matrix with the hybrid approach

flights in this subset. If a deadhead flight (dummy row) is included in any column, the crew deadhead cost is added to this trippair. The dummy rows that do not have to be covered are not provided as part of the coverage constraints in the set partitioning formulation. The dummy deadhead rows are dropped from the problem just before they are provided to the optimizer.

Crew Rostering

The Business Process of Crew Rostering

For each crew member, a monthly LOF is constructed out of trippairs. The LOF considers the crew vacation and training requirements over the month. The crew members are then assigned to these monthly lines either by auctions or on tha basis of seniority. There are two different approaches in which air carriers implement crew rostering. For example, most air carriers in North America use the bidlines approach, where anonymous rosters (or so-called bidlines) are first created. Each roster is composed of a number of trippairs, vacation activities, standby and reserve activities, and training activities. Then, the individual crew members bid on these anonymous rosters, based on their seniority. Each crew member selects the roster that matches their preferences regarding work and vacation duties during the month.

Another rostering approach, which is adopted by most European air carriers, is called personalized rostering. In this approach, individual rosters are constructed directly for each crew member. The roster is created such that it satisfies a predefined set of preferences that has been expressed by each crew member. Sometimes, an individual's preferences may conflict with other crew members' preferences, so a fair rule should be applied to guarantee that the generated rosters satisfy a minimum list of preferences for each crew member. Next, the additional preferences and conflicting preferences are awarded on the basis of crew seniority, with the most senior crew members getting the maximum of their predefined preferences. Earlier versions of this approach only considered fair distribution of work rosters, with no preferences given to the senior crew.

In the bidlines approach, individual crew members know exactly the roster assigned to them because, by bidding for a specific line, they know exactly what they will get if the bid is granted. In contrast, in the personalized rostering approach, crew members do not know what they will get when they define their preferences. They only express preferences for certain attributes of their rosters, and they have to wait until their roster is constructed. A main drawback of the bidlines approach is that the planning process consists of more steps, because individual crew members have to bid on the constructed rosters. In contrast, in the personalized rostering approach, individual rosters are assigned directly to each crew member once they are generated.

Rules and Regulations in the Rostering Process

Several rules and regulations should be considered when generating crew rosters. These rules are imposed by the aviation governmental agencies (for example, the FAA in the US), by the air carrier itself, and by agreements between the air carrier and crew unions. According to Khol and Karrisch (2004), these rules are classified as horizontal rules, vertical rules, and artificial rules. Since rosters are presented as rows, the rules that depend on one roster and a crew member are called horizontal rules, while the vertical rules include information across several rows (several rosters and crew members). The artificial rules represent additional constraints that are used to exclude the (feasible) solutions of poor quality.

Horizontal rules Khol and Karrisch (2004) indicated that large European air carriers usually have more than 100 horizontal rules and regulations to consider for crew rostering. They presented a few examples of the most important horizontal rules. The first rule is to guarantee compatibility between the crew member and the tasks included in the roster. Another rule is to ensure the sequence of tasks in the roster, which may include the rest time between tasks. Another rule is to consider the cumulated activities, such as the total flying or block hours over a given period or the total rest period within a given limit. The constraint on the accumulated value may be imposed in fixed intervals or gliding intervals. Fixed intervals include calendar month or calendar week, while gliding intervals include all seven consecutive calendar days or all 168 consecutive hour periods. The rule may be applied to a working period regardless of its length. Other rules might be imposed to limit tiring traveling across several time zones. Finally, the roster should consider the history of crew member activities in the previous month.

Vertical rules The vertical rules consider more than one roster at a time and may depend on a subset of rosters or the whole schedule. The basic vertical rules are crew complement, qualification-type, and global constraints. Usually, most of the different crew activities require different crew complements. For example, for a short-haul cockpit crew problem, the flight pairings are usually assigned to one captain and one F/O. The qualification-type constraints are for the tasks that require some crew members to hold particular qualifications. This constraint is mainly a cabin problem, but it also occurs for the cockpit crew. These constraints are applied to the number of inexperienced crew on a task. The incompatibilities constraints are used to avoid assigning incompatible crew to the same task. The language qualification for flight attendants on international flights is another example of qualification-type constraints. The global constraints put requirements on the entire solution. Examples of global constraints include the upper bound on the costs of the solution: for example, one may wish to maximize the crew satisfaction subject to a constraint on the total costs. The global constraints also take care of training programs, where the total number of training instructors has to be distributed among the different generated rosters (Khol and Karrisch 2004).

Artificial rules Artificial rules (also called quality rules) are based mainly on the experience of the planners and provide a further restriction on the solution space such that unattractive solutions are omitted. Artificial rules are considered to improve the robustness of the solution or support for the solution methods. For example, to obtain a better performance from the solution methods, some bad rosters may be excluded from the generation process. These excluded rosters are unlikely to be a part of a good solution.

Crew Rostering Cost

The main objective of the crew rostering problem is to cover all the available crew activities with the minimum possible cost and with a robust schedule. Another important objective is to improve the crew's quality of life by satisfying the maximum number of their preferences. The major cost of the rostering problem is incurred when a crew member task is not covered or an activity is unassigned. However, the existence of an unassigned activity does not necessarily mean that the solution is not feasible. Even if the unassigned activity is a pairing, it may be possible to solve the problem by asking the crew to do overtime, by reducing the number of reserves, by hiring new crew, or by canceling the flight (Khol and Karrisch, 2004).

Typically, the cost of having an open task can be calculated *a priori*. This cost is a function of the type of task and its length. The crew members are often paid a regular salary for up to a certain number of block hours per year. The block hours that exceed this amount are paid separately. In this situation, scheduling should avoid situations where some crew members receive overtime payments, while others fly fewer hours than they are permitted to fly. Other important cost components are usually related to training assignments (for example, simulator training). If training is away from base, positioning the flights (also called deadhead flights), hotel costs, and so on all need to be considered. Moving pairings between bases requires deadheads for crew. The crew members who are temporarily moved to another base may receive extra compensation. An additional cost may be added to any roster that is expected to cause problems in operation. Other costs may be added to generate rosters with particular attributes that are related to the total number of block hours, duty hours, days off, and the number of days away from home. Equalization of the workload among the crew is a very significant issue, which should be considered within the planning period. To implement the equal assignment criteria, it is necessary to define the penalties for deviations from the average values. This function is nonlinear in general, because a few large deviations may be worse than many small deviations.

The Crew Rostering Problem Formulation and Solution Methodology

The crew rostering problem is formulated in a similar way to the crew pairing problem, with flights in the crew pairing problem being replaced by crew activities in the crew rostering problem. The generate-and-optimize principle is also usually used to solve the crew rostering problems, similar to the technique

applied in solving the crew pairing problem. The generate-and-optimize-principle solves the problem in two main steps. In the first step, which is known as the generation problem, a large set of feasible monthly rosters is created, then another optimization problem is solved. In the optimization problem, a set partitioning-type problem is solved to select one roster for each crew member. The main objective is to minimize cost, satisfying both the demands of the activities and the constraints within and across the crew members. Since it is not possible to have an explicit representation of all possible rosters, the master problem is always defined by a subset of all possible rosters. In the (generation) sub-problem, a large number of legal rosters is generated. This can be done by partial enumeration-based techniques that are constrained to the shortest path of the column generation (Khol and Karrisch 2004).

Primary Contributions

Crew Planning Literature

Rubin (1973) made an early contribution to the airlines' crew scheduling problem. In this work, a program called 'TPACS' is developed by IBM to improve the manual solutions that are generated by human schedulers at United Airlines. TPACS uses a local optimization to improve the initial solution. It takes the solution pairings two or three at a time and tries to reorganize the flight segments in new pairings that have a better performance than the original pairings. The program replaces the original solution pairings in the solution with the new pairings and proceeds in this manner until all combinations are tried. These early efforts found that a global optimization to the problem is difficult to achieve. Subsequent research focuses on the development of efficient heuristic procedures to address the problem.

American Airlines and Continental Airlines use a system called TRIP (Gershkoff 1989). The primary objective is to minimize the cost of flying the published airline schedule, given the operational crew constraints. In this system, an optimization model that uses a set partitioning framework is used. In this formulation, the rows represent flights to be covered, and the columns represent candidate crew trips. The crew scheduling problem is formulated as an integer programming problem and can be solved using commercial optimization software. To solve this problem, the following algorithm is considered. First, an initial feasible solution is generated (either manually or by using an interactive computer system). Next, TRIP randomly selects a set of pairings from the current best solution. The selected pairings give a disjointed solution to the set partitioning problem. The flight segments in the selected pairings are considered as a sub-problem. All possible pairings are generated for the sub-problem, and the resulting set partitioning problem is solved. If an improved solution is found, it replaces the selected set of pairings to become part of the current best solution. TRIP then selects another set of pairings randomly and repeats the process. Anbil et al. (1991) and Anbil et al. (1992) discuss some

of the operational benefits of implementing such a decision support system and its overall impact on the finances of American Airlines.

A new crew scheduling optimization system developed for United Airlines (Graves et al. 1993) allows a fast response to schedule changes and reduces crew scheduling costs. The system has two main components, a generator and an optimizer. The generator creates candidate crew trips which are fed as variables to the optimizer as an elastic set partitioning integer programming problem. The optimizer then seeks to find a set of pairings that covers all of the flight segments exactly once at a minimal cost. The system continues to iterate between the generator and optimizer to improve the solution. It is claimed that this system generates savings of $16,000,000 annually in crew scheduling costs.

It is well known that airlines' ability to solve the crew scheduling problem depends on the size of the constraint matrix in the mathematical formulation of the problem. The size of the constraint matrix is proportional to the number of flights to be covered in the schedule. Recently, significant research efforts have been focused on developing solution techniques to help in the timely solution of such large-scale crew scheduling problems. Initially, attention was given to a hybrid technique called the branch-and-cut procedure (Hoffman and Padberg 1993). In this method, classes of valid inequalities are left out of the linear predictive relaxation algorithm. Many constraints exist to handle the inequalities efficiently, and most of them will not be binding in an optimal solution anyway. If an optimal solution to an LP relaxation is infeasible, a sub-problem is solved to try to identify the violated inequalities in a class. If one or more of the violated inequalities are found, some are added to the LP to cut off the infeasible solution. Then, the LP is re-optimized. Branching occurs when no violated inequalities are found to cut off an infeasible solution.

In the last decade, attention has been given to another solution procedure known as the branch-and-price technique, which combines the standard Internet Protocol (IP) solution technique of branch-and-bound with the column generation. A recent review of a major contribution to this technique has been undertaken by Barnhart and Shenoi (1996), Barnhart et al. (1998), and Barnhart et al. (2003). The main idea of branch-and-price is similar to that of branch-and-cut. The difference is that the procedure focuses on column generation rather than row generation. In branch-and-price, columns are left out of the LP relaxation because there are too many columns to handle the columns efficiently. To validate the optimality of a LP solution a sub-problem called the pricing problem is solved to try to determine the columns to enter the basis. If such columns are found, the LP is re-optimized. Branching occurs when no columns price out to enter the basis, and the LP solution does not satisfy the integrality conditions. In branch-and-price-type algorithms there are three approaches to solve the sub-problem (Klabjan 2006). The first approach is the constrained shortest path (Desrosiers et al. 1991, Lavoie et al. 1988, Anbil et al. 1994, and Vance et al. 1997). Another approach is based on solving the k-shortest path algorithm, which is used by Carmen Systems (Galia and

Hjorring 2003). The third approach is to perform a depth-first search enumeration of pairings on a network, (Klabjan et al. 2001). As for the branching rule used in the branch-and-price algorithms, the most widely used rule is to branch on the follow-ons. This branching rule is used in Desrosiers et al. (1991), Vance et al. (1997), and Anbil et al. (1994). Other literature for solving a large-scale crew pairing problem include Anderson et al. (1998), Chu et al. (1997), Grönkvist (1998), Gustafsson (1999), Vance et al (1995), Wedlin (1994), Desaulniers et al. (1997), Mingozzi et al. (1999), Klabjan et al. (2001), and Klabjan et al. (2002).

Crew Rostering Literature

For the bidlines approach, Campbell et al. (1997) describe the crew rostering application at FedEx and Christou et al. (1999) describe the system deployed at Delta Air Lines. In these two systems, the problem is solved using meta-heuristics. Jarrah and Diamond (1997) develop a semi-automatic system that is based on column generation. In this system, the user has the ability to influence the subset of columns to be generated.

For personalized rostering, the most commonly used approach is the set partitioning model, as described in the literature. For example, The GERAD research center developed a column generation approach which the sub-problems that generate the rosters for individuals are solved as the constrained shortest paths. This approach is used within a commercial system (AD OPT ALTITUDE). A description of the application of this system is given by Gamache and Soumis (1998), Gamache et al. (1998), Gamache et al. (1999), and Lasry et al. (2000).

Day and Ryan (1997) describe the rostering approach of flight attendants for Air New Zealand's short-haul operations. The problem is decomposed into the problem of assigning off-days followed by the problem of assigning the pairings and other activities. The column generation model is used to solve both problems. The advantage of this approach is that the total number of columns in each of the two problems is limited compared to the case where the off-day assignment and pairing assignment is optimized simultaneously.

Cappanera and Gallo (2001) present an exact IP for crew rostering. The model is applied to solve small instances from a medium-sized Italian carrier. More recently, Cappanera and Gallo (2004) formulated the crew rostering problem by using the multi-commodity flow approach. El Moudani et al. (2001) suggest a heuristic approach to solving a bi-criteria version of the rostering problem to be applied to medium-sized problems. The two criteria are related to crew satisfaction and cost.

A solution approach, combining the column generation with constraint programming (CP) has been developed within the Parrot project (Fahle et al. 2002, Sellmann et al. 2002). CP is used to trim the search for rosters, whereas the master problem that selects a roster for all crew members is solved as a linear program. Dawid et al. (2001) and König and Strauss (2000a, 2000b) describe an

enumeration heuristic with propagation techniques. The heuristic is implemented in the SWIFTROSTER algorithm and applied to data from a medium-sized European airline. The Carmen Crew Rostering system is described in Augustsson (1997), Hjorring et al. (1999), Kohl (1999), Kohl and Karisch (2000), and more recently in Kohl and Karisch (2004). The Carmen Crew Rostering system is currently in use at several major European airlines including British Airways, KLM, Iberia, Alitalia, and Scandinavian Airlines (SAS), as well as at one of the world's largest passenger transportation companies, Deutsche Bahn (German State Railways).

References

Anbil, R., Gelman, E., Patty, B., and Tanga, R. 1991. Recent Advances in Crew-Pairing Optimization at American Airlines. *Interfaces* 21(1), 62-74.

Anbil, R., Tanga, R., and Johnson, E. 1992. A Global Approach to Crew-pairing Optimization. *IBM Systems Journal*, 31, 71-78.

Anbil, R., Barnhart, C., Johnson, E., and Hatay, L. 1994. A Column Generation Technique for the Long-Haul Crew Assignment Problem. In T.A. Ciriani and R.C. Leachman (eds), *Optimization in Industry II*. John Wiley & Sons, 7-24.

Andersson, E., Housos, E., Kohl, N., and Wedelin, D. 1998. Crew Pairing Optimization. In G. Yu (ed.), *Operations Research in the Airline Industry*. Kluwer Academic Publishers, Boston, MA.

Augustsson, L. 1997. Partial Evaluation in Aircraft Crew Planning. *ACM SIGPLAN Notices*, 23(12), 127-136.

Barnhart, C. and Shenoi, R. 1996. A Column Generation Technique for the Long-Haul Crew Assignment Problem. In T.A. Cirani and R.C. Leachman (eds), *Optimization in Industry II.* Wiley, New York.

Barnhart, C., Cohn, A., Johnson, E., Klabjan, D., Nemhauser, G., and Vance, P. 2003. Airline Crew Scheduling. In R.W. Hall, (ed.), *Handbook of Transportation Science*. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Barnhart, C., Johnson, E.L., Nemhauser, G.L., Savelsbergh, M.W.P., and Vance, P.H. 1998. Branch-and-Price: Column Generation for Solving Huge Integer Programs. *Operations Research*, 46, 316-329.

Campbell, K.W., Durfee, R.B., and Hines G.S. (1997). FedEx Generates Bid Lines Using Simulated Annealing. *Interfaces*, 27(2), 1-16.

Cappanera, P. and Gallo. G. 2001. *On the Airline Crew Rostering Problem*. Technical Report TR-01-08, Department of Computer Science, University of Pisa, Italy.

Cappanera, P. and Gallo. G. 2004. A Multicommodity Flow Approach to the Crew Rostering Problem. *Operations Research*, 52(4), 583-596.

Christou, I.T., Zakarian, A. Liu, J., and Carter. H. 1999. A Two-Phase Genetic Algorithm for Large-Scale Bidline-Generation Problems at Delta Air Lines. *Interfaces*, 29(5), 51-65.

Chu, H.D., Gelman, E., and Johnson, E.L. 1997. Solving Large Scale Crew Scheduling Problems. *European Journal of Operational Research*, 97, 260-268.

Dawid, H., König, J., and Strauss, C. 2001. An Enhanced Rostering Model for Airline Crews. *Computers and Operations Research*, 28(7), 671-688.

Day, P.R. and Ryan. D.M. 1997. Flight Attendant Rostering for Short-Haul Airline Operations. *Operations Research*, 45(5), 649-661.

Desaulniers, G., Desrosiers, J., Dumas, Y., Marc, S., Rioux, B., Solomon, M.M., and Soumis, F. 1997. Crew Pairing at Air France. *European Journal of Operational Research*, 97, 245-259.

Desrosiers, J., Dumas, Y., Desrochers, M., Soumis, F., Sanso, B., and Trudeau, P. 1991. *A Breakthrough in Airline Crew Scheduling*. Technical Report G-91-11, GERAD.

El Moudani, W., Nunes Cosenza, C.A., de Coligny, M., and Mora-Camino F. 2001. A Bi-Criterion Approach for the Airlines Crew Rostering Problem. In E. Zitzler, K. Deb, L. Thiele, C.A. Coello, and D. Corne (eds), *First International Conference on Evolutionary Multi-Criterion Optimization*. Lecture Notes in Computer Science, Vol. 1993. Springer, Berlin, pp. 486-500.

Fahle, T., Junker, U., Karisch, S.E., Kohl, N., Sellmann, M., and Vaaben, B. 2002. Constraint Programming Based Column Generation for Crew Assignment. *Journal of Heuristics*, 8(1), 59-81.

Galia, R. and Hjorring, C. 2003. *Modeling of Complex Costs and Rules in a Crew Pairing Column Generator*. Technical Report CRTR-0304, Carmen Systems, Sweden.

Gamache, M. and F. Soumis. 1998. A Method for Optimally Solving the Rostering Problem. In G. Yu (ed.), *Operations Research in the Airline Industry*. Kluwer, Norwell, MA, pp. 124-157

Gamache, F., Soumis, F., Villeneuve, D., Desrosiers, J., and Gélinas, E. 1998. The Preferential Bidding System at Air Canada. *Transportation Science*, 32(3), 246-255.

Gamache, M., Soumis, F., Marquis, G., and Desrosiers, J. 1999. A Column Generation Approach for Large Scale Aircrew Rostering Problems. *Operations Research*, 47(2), 247-262.

Gershkoff, I. 1989. Optimizing Flight Crew Schedules. *Interfaces*, 19, 29-43.

Graves, G.W., McBride, R.D., Gershkoff, I., Anderson, D., and Mahidhara, D. 1993. Flight Crew Scheduling. *Management Science*, 39, 736-745.

Grönkvist, M. 1998. *Structure in Airline Crew Optimization Problems*. Master's thesis, Chalmers University of Technology, Gothenburg, Sweden.

Gustafsson, T. 1999. *A Heuristic Approach to Column Generation for Airline Crew Scheduling*. Lic. thesis, Chalmers University of Technology, Gothenburg, Sweden.

Hjorring, C.A., Karisch, S.E., and Kohl, N. 1999. Carmen Systems' Recent Advances in Crew Scheduling. In *Proceedings of the 39th Annual AGIFORS Symposium*, New Orleans, USA, October 3-8, 404-420.

Hoffman, K.L. and Padberg, M. 1993. Solving Airline Crew Scheduling Problems by Branch-and-Cut. *Management Science*, 39(6), 657-682.

Jarrah, A.I.Z. and Diamond, J.T. 1997. The Problem of Generating Crew Bidlines. *Interfaces*, 27(4), 49-64.

Klabjan, D. 2006. Large-Scale Models in the Airline Industry. In G. Desaulniers, J. Desrosiers, and M.M Solomon, (eds), *Column Generation*, Springer, New York, 163-195.

Klabjan, D., Johnson, E.L., Nemhauser, G.L., Gelman, E., and Ramaswamy, S. 2001. Solving Large Airline Crew Scheduling Problems: Random Pairing Generation and Strong Branching. *Computational Optimization and Applications*, 20(1), 73-91.

Klabjan, D., Johnson, E.L., Nemhauser, G.L., Gelman, E., and Ramaswamy, S. 2002. Airline Crew Scheduling with Time Windows and Plane-Count Constraints. *Transportation Science*, 36(3), 337-348.

Kohl, N. 1999. The Use of Linear and Integer Programming in Airline Crew Scheduling. In *Proceedings of the 3rd Scandinavian Workshop on Linear Programming*, Lyngby, Denmark.

Kohl, N. and Karisch, S.E. 2000. Integrating Operations Research and Constraint Programming Techniques in Crew Scheduling. In *Proceedings of the 40th Annual AGIFORS Symposium,* Istanbul, Turkey.

Kohl, N. and Karisch, S.E. 2004. Airline Crew Rostering: Problem Types, Modeling and Optimization. *Annals of Operations Research*, 127, 223-257.

König, H. and Strauss, C. 2000a. Rostering-Integrated Services and Crew Efficiency. *Information Technology and Tourism*, 3(1), 27-39.

König, H. and Strauss, C. 2000b. Supplements in Airline Cabin Service. In D. Buhalis, D.R. Fesenmaier and S. Klein (eds), *Information and Communication Technologies in Tourism 2000*. Springer, Berlin, pp. 365-374.

Lasry, A., McInnis, D., Soumis, F., Desrosiers, J., and Solomon, M.M. 2000. Air Transat Uses ALTITUDE to Manage Its Aircraft Routing, Crew Pairing, and Work Assignment. *Interfaces*, 30(2), 41-53.

Lavoie, S., Minoux, M., and Odier, E. 1988. A New Approach for Crew Pairing Problems by Column Generation with an Application to Air Transport. *European Journal of Operational Research*, 35, 45-58.

Mingozzi, A., Boschetti, M.A., Ricciardelli, S., and Bianco, L. 1999. A Set Partitioning Approach to the Crew Scheduling Problem. *Operations Research*, 47(6), 873-888.

Rubin, J. 1973. A Technique for the Solution of Massive Set Covering Problems, with Applications to Airline Crew Scheduling. *Transportation Science*, 7, 34-48.

Sellmann, M., Zervoudakis, K. Stamatopoulos, P., and Fahle, T. 2002. Crew Assignment via Constraint Programming: Integrating Column Generation and Heuristic Tree Search. *Annals of Operations Research*, 115, 207-225.

Vance, P., Atamtürk, A., Barnhart, C., Gelman, E., Johnson, E., Krishna, A., Mahidhara, D., Nemhauser, G., and Rebello, R. 1997. *A Heuristic Branch-*

and-Price Approach for the Airline Crew Pairing Problem. Technical Report LEC-97-06, Georgia Institute of Technology.

Vance, P., Barnhart, C., Johnson, E.L., and Nemhauser, G.L. 1995. Airline Crew Scheduling: A New Formulation and Decomposition Algorithm. *Operations Research*, 45, 188-200.

Wedelin, D. 1994. An Algorithm for Large Scale 0-1 Integer Programming with Application to Airline Crew Scheduling. *Annals of Operations Research*, 57, 283-301.

Chapter 7
Gate Assignment

Introduction

The problem of the aircraft-gate assignment was defined with the introduction of gate walkways that connect passengers directly between the aircraft and the terminal area. Accordingly, each aircraft has to be allocated to one gate at the terminal. In the case that not enough gates are available, aircraft are parked at the apron and passengers are serviced by buses to connect to the terminal. The gate assignment problem is more significant at busy airports (hubs), where there are many arrivals and departures scheduled within a short time period over a limited number of gates. As mentioned earlier, at hub stations, the air carriers schedule time banks where a few flight arrivals are followed by a few flight departures to allow passengers to connect between different origins and destinations.

The gate assignment at the hub affects the walking distance of connecting passengers in the terminal, the baggage handling of connecting passengers, and cargo transfer, which directly impacts on the air carrier's operating cost. For example, Figure 7.1 shows the trajectories of the connecting passengers and baggage of one flight arrival at a hypothetical airport terminal. It is clear that the total travel distance of the connecting passengers and the total transportation distance of connecting baggage depend on the location of the gate assigned to this flight. Proper assignment of the gates for each flight can optimize the connection time of passengers and baggage, and improve the overall turn time of aircraft at the terminal. This optimal gate assignment can contribute to the productivity of the aircraft and overall revenue to the airline.

There are several conflicting objectives for the gate assignment problem. First, gate assignment should increase the convenience of connecting passengers by minimizing the total walking distance. Furthermore, gate assignment should consider minimizing the total operating cost related to baggage and cargo handling, and the aircraft taxi distance. In addition, it should consider matching the aircraft size to the appropriate gate to account for the passengers' seating capacity at the gate. Another consideration is aircraft assignment at adjacent gates: for example, two aircraft with long wings cannot occupy two adjacent gates at the same time. A further factor to be taken into account is schedule variability due to irregular operation conditions. So gate assignment should be sufficiently robust to minimize changes during irregularities.

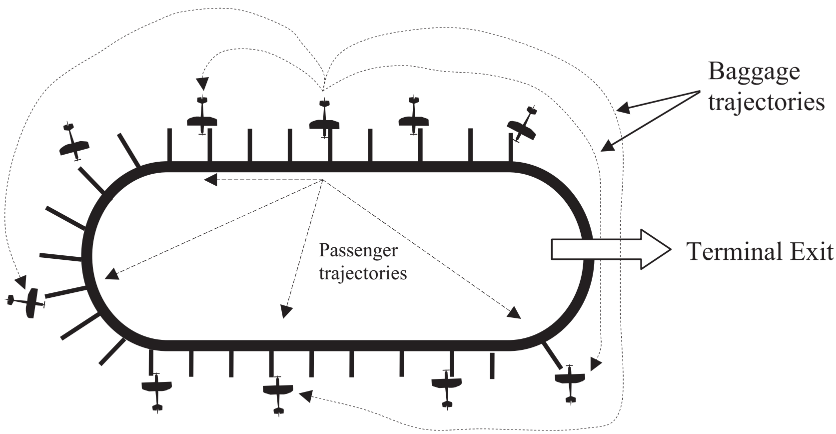


Figure 7.1 Example of trajectories of connecting passengers/baggage at airport terminal

The gate assignment problem is solved during the planning and operation phases of the schedule. At the planning level gate assignment should ensure that all flights can be accommodated by the gates available to the air carrier. During the operation phase, flights might need to be reassigned to other gates as a result of schedule irregularities due to flight delays and cancellations.

Problem Characteristics

Figure 7.2 gives an example of the aircraft assignment for a hypothetical set of gates. As shown in the figure, the solution to the gate assignment problem for each gate consists of a set of flight service times (aircraft turns) separated by periods of idle time at the gate. The turn time of an aircraft represents the difference between the departure time of a flight and the arrival time of its preceding flight at the gate. The idle time at the gate starts at the departure time of a flight and ends at the arrival time of its next flight at the gate.

The gate assignment problem is characterized as a non-deterministic problem and involves significant uncertainties that are related to the flight delays and cancellations. A delay to an outbound flight cause the aircraft to occupy the gate for a longer period than planned for this flight. Similarly, a delay to an inbound flight will also disrupt the assignment schedule at the gate. Furthermore, a cancellation of an outbound flight generates a need for an additional aircraft to be parked at one of the gates or near the terminal area, until a new assignment is given to this aircraft.

The gate assignment problem is also characterized as being time-critical during operation. During irregular operation conditions, a reassignment (recovery) plan needs to be rapidly formulated and disseminated quickly to the

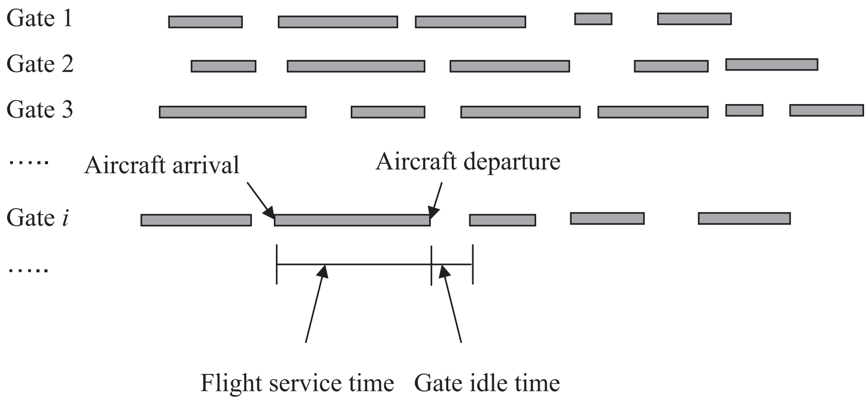


Figure 7.2 Example of the flight assignment for a hypothetical set of gates

operators so that the passengers are informed about the gate changes, if any, in a timely manner.

Finally, the gate assignment problem is known to be combinatorial in nature, with many interacting decisions. The assignment of one flight to a particular gate affects the decision on which flight is to follow at this gate. It also affects the decision on which flight can be assigned to the adjacent gates, due to wing-length constraints for the corresponding aircraft.

Problem Formulation

The gate assignment problem is formulated by an objective function and a set of constraints. The objective function involves minimizing the number of ungated flights, the total walking distance of passengers, the baggage and cargo handling cost, and the aircraft taxi time. The constraints of the problem are related to the number of available gates that are suitable for each aircraft type, the flight timing constraints (arrivals and departures), and the constraints from adjacent gates.

Most commercial air carriers follow heuristic approaches to solve the gate assignment problem. These models vary significantly in their capabilities and solution quality. Here, we present an example of these heuristics offered by Ding et al. (2005a, 2005b). They considered the case of an overconstrained gate assignment, where the number of flights exceeds the number of gates during some periods of the day. The solution methodology considered minimizing the number of ungated flights and the total walking distance of passengers.

Consider that:

n = the set of flights arriving at the airport
m = the set of gates available at the airport

$m+1$ = index of the apron. The 0 index refers to the airport entrance (check-in) or exit
a_i = the arrival time of the flight i (that is, the start of the aircraft turn)
d_i = the departure time of the flight i (that is, the end of the aircraft turn)
$w_{k,l}$ = walking distance of passengers from gate k to gate l
$f_{i,j}$ = number of passengers transferring from flight i to flight j
$f_{0,i}$ = number of passengers starting a trip from gate i
$f_{i,0}$ = number of passengers ending a trip from gate i.

The decision variables are defined as follows:

$$y_{i,k} = \begin{cases} 1, \text{ if flight } i \text{ is assigned to gate } k, \ (0 < k \leq m+1) \\ 0, \text{ otherwise} \end{cases}$$

One critical set of constraints should be satisfied in order to guarantee that any two flights i and j do not overlap when assigned to the same gate.

For any gate k ($k \neq m+1$) and for any two flights i and j, $y_{i,k} = y_{j,k} = 1$, if $a_i > d_j$ or $a_j > d_i$. For any gate (not including the apron area), any two flights can be assigned to the same gate if they do not occupy the gate at the same time.

The problem is formulated mathematically in a quadratic form as follows:

The objective function is:

$$\text{minimize} \sum_{i=1}^{n} y_{i,m+1}$$

$$\text{minimize} \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{m+1}\sum_{l=1}^{m+1} f_{i,j} w_{k,l} y_{i,k} y_{j,l} + \sum_{i=1}^{n} f_{0,i} w_{0,i} + \sum_{i=1}^{n} f_{i,0} w_{i,0} \tag{7.1}$$

Subject to:

$$\sum_{k=1}^{m+1} y_{i,k} = 1, \ (\forall i, 1 \leq i \leq n) \tag{7.2}$$

$$a_i < d_i, \ (\forall i, 1 \leq i \leq n) \tag{7.3}$$

$$\begin{gathered} y_{i,k} y_{j,k} (d_j - a_i)(d_i - a_j) \leq 0, \\ (\forall i, \ j, \ 1 \leq i \leq n, \ 1 \leq j \leq n, \ \forall k, \ \ 1 \leq k \leq m \ \ (k \neq m+1)) \end{gathered} \tag{7.4}$$

$$y_{i,k} \in \{0,1\}, \ (\forall i, 1 \leq i \leq n, \forall k, 1 \leq k \leq m+1) \tag{7.5}$$

The objective function given in equation 7.1 is to minimize both the number of flights assigned to the apron and the total walking distance of connecting passengers as well as passengers that are starting and ending their trip at the terminal. The first set of constraints given in equation 7.2 is to ensure that each

flight is assigned to one and only one gate, or assigned to the apron. The second set of constraints given in equation 7.3 specifies that, for each flight turn, the departure time is later than the arrival time. The next set of constraints is to ensure that two flights cannot overlap if assigned to the same gate (excluding the apron), as explained above. The final set of constraints is to ensure that the decision variables are binary.

Solution Methodology

The mathematical formulation presented above is characterized as nonlinear and, accordingly, its solution is intractable if available mathematical programming solvers are used. Thus, a heuristic approach is developed to find a solution for the gate assignment problem. The proposed heuristics consist of two main steps as presented by Ding et al. (2005a, 2005b). First, a greedy algorithm is used to find an initial feasible solution, in which the number of ungated flights is minimized. This greedy approach follows the solution of a widely known operations research problem known as the activity selection problem (ASP) (Cormen et al. 2001). Second, the solution is improved iteratively using a Tabu-search to minimize the total walking distance. Tabu-search is an iterative procedure designed for the solution of optimization problems and belongs to the class of local search techniques used to solve optimization problems by tracking and guiding the search. Tabu-search enhances the performance of a local search method by marking a potential solution once it has been determined, so that the algorithm does not visit that possibility repeatedly (Glover and Laguna 1997).

The main steps of the greedy algorithm can be described as follows:

1. Sort the flights according to the departure time (the end of the aircraft turn) d_i, $(1 \leq i \leq n)$.
   Let g_k $(1 \leq k \leq m)$ represent the departure time of last flight on gate k.
   Set $g_k = -1$ for all k.
2. For each flight i:
   (a) Find gate k with the maximum g_k such that $g_k < a_i$.
   (b) If such k exists, assign flight i to gate k and update $g_k = a_i$.
   (c) If k does not exist, assign flight i to the apron.
3. Output the results.

The algorithm is applied to the example of the eight flights shown in Figure 7.3. As shown in the figure, the aircraft turns are sorted according to their departure time (that is, at the end of the aircraft turn). Turn 1 has the first departure, and turn 8 has the last departure. At the outset, all gates are empty. Therefore, we set $g_k = -1$ for all gates. Then, we assign turn 1 to gate 1 and update $g_1 = a_1$. We then assign turn 2 to gate 2 and update $g_2 = a_2$. Turn 3 is assigned to gate 3. Turn 4 is

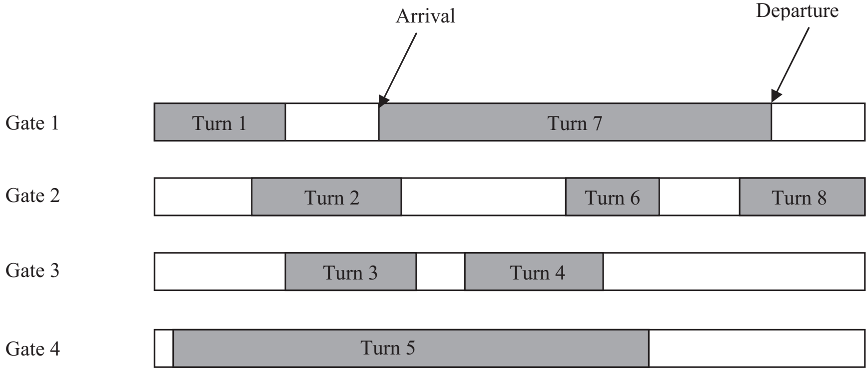


Figure 7.3 Example of eight flights assigned to gates

also assigned to gate 3 because it is the latest available gate before the start of turn 4. We continue until all aircraft turns are assigned to the different gates. If no gate is available to a flight, we assume that the aircraft is assigned to the apron.

Up to this point, and when assigning flights to the different gates, no attention is given to evaluating the walking distance of connecting passengers between gates in the terminal. Thus, we now use the Tabu-search methodology to improve the solution given by the greedy algorithm in order to minimize passengers' total walking distance. There are three main neighborhood search methods considered in this methodology. The first method is denoted as the insert move, which involves a move of a single aircraft turn to a gate other than the one currently assigned. The second is the interval exchange move that involves the exchange of two intervals in the current assignment. (An interval consists of one or more consecutive turns at one gate.) Finally, the third method is the apron exchange move, which involves the exchange of one turn that is assigned to the apron with a turn that is assigned currently to the gate. The Tabu-search methodology continues until the solution is no longer improving.

Primary Contributions

In the past few decades, many optimization methods have been reported to improve the gate assignment operation at airport terminals. An early contribution to the gate assignment problem is that of Baron (1969), who presents a simulation model of the airport operation. Braaksma and Shortreed (1971) provide one of the first attempts to model the assignment of aircraft to gates, while minimizing travel distances. The total passenger walking distance is based on the passenger embarkation and disembarkation volumes, transfer passenger volumes, gate-to-gate distances, check-in-to-gate distances, and aircraft-to-gate assignments. The cost associated with the placing of an aircraft at a gate depends on the distances

from key facilities as well as the relations between these facilities. Obata (1979) presents an evaluation of exact and heuristic algorithms for the gate assignment problem and shows that it is a quadratic assignment problem and NP-hard. Babic et al. (1984) formulate the problem as a 0-1 integer program and use branch-and-bound algorithm to find an optimal solution. The main limitation is that the transfer passengers are not considered in the model. Transfer passengers are, however, considered in the work of Mangoubi and Mathaisel (1985). They use a LP relaxation of Babic et al.'s (1984) integer program, with greedy heuristics, to solve the problem.

Bihr (1990) presents a simplified conceptual solution to the aircraft gate assignment problem using 0-1 linear programming and minimizing the walking distance problem for the fixed arrivals in a hub operation. Wirasinghe and Bandara (1990) consider the cost of delays in an approximate closed-form solution. Bandara and Wirasinghe (1992) develop a model for minimizing walking distance for airport terminal configurations. Yan and Chang (1998) propose a network model for gate assignment, in which the problem is formulated as a multi-commodity network flow problem. Bolat (1999) develops a mathematical model to assign the flights with the minimum range of unutilized time periods of gates, subject to the level of service offered to passengers and other physical and managerial considerations. More recently, a model is considered by Ruperto and Francesc (2007) that minimizes the total taxiing time in the system for both passengers and aircraft.

Robustness and punctuality are also considered in the gate assignment problem. For example, Hassounah and Steuart (1993), plan scheduled buffer times to improve punctuality. Yan and Chang (1998) and Yan and Huo (2001) add a fixed buffer time between flights assigned to a gate. Gate assignment with time windows is studied by Zhu et al. (2003). Also, Bolat (2000) provides procedures for robust gate assignments for arriving aircraft.

Because the gate assignment problem is NP-hard, researchers use simulation and heuristic approaches. For example, in addition to the work of Baron (1969), simulation-based models for the gate assignment problem are also considered by Cheng (1998a, 1998b). Haghani and Chen (1998) propose a heuristic that assigns successive flights to the same gate when there is no overlapping. Where overlapping does occur, flights are assigned on the basis of the shortest walking distance coefficients. In Lam et al. (2002), an intelligent agent for the airport gate assignment is developed to capture schedule dynamics and gate changes at an airport. The agent aims to perform the gate assignment for every flight, taking into consideration the gate and flight dynamics, transfers, requirements of the airlines, aircraft types, airport operation rules, and so on. A knowledge-based expert system forms the core of the system and this is connected to external databases for flight and passenger information. Real-time changes to airport gates and flights can be made through a graphical user interface with the capability of performing real-time updates of the results and information. The simulation annealing approach is considered by Kirkpatrick et al. (1983). Expert systems for the gate assignment

problem are also considered by Gosling (1990), and Srihari and Muthukrishnan (1991). Genetic algorithm approaches for the gate assignment problem are considered by Gu and Chung (1999), Bolat (2001), and Hu and Chen (2005). Tabu-search approaches are considered by Xu and Bailey (2001), and Ding et al. (2005a, 2005b).

References

Babic, O., Teodorovic D., and Tosic V. 1984. Aircraft Stand Assignment to Minimize Walking. *Journal of Transportation Engineering*, 110, 55-66.

Bandara, S., and Wirasinghe, S.C. 1992. Walking Distance Minimization for Airport Terminal Configurations. *Transportation Research* A, 26(1), 59-74.

Baron, P. 1969. A Simulation Analysis of Airport Terminal Operations. *Transportation Research*, 3(3), 481-491.

Bihr, R. 1990. A Conceptual Solution to the Aircraft Gate Assignment Problem using 0,1 Linear Programming. *Computers and Industrial Engineering*, 19(1-4), 280-284.

Bolat, A. 1999. Assigning Arriving Aircraft Flights at an Airport to Available Gates. *Journal of the Operational Research Society*, 50, 23-34.

Bolat, A. 2000. Procedures for Providing Robust Gate Assignments for Arriving Aircraft. *European Journal of Operations Research*, 120, 63-80.

Bolat, A. 2001. Models and a Genetic Algorithm for the Static Aircraft-Gate Assignment Problem. *Journal of the Operational Research Society*, 52(10), 1107-1120.

Braaksma, J.P., and Shortreed, J.H. 1971. Improving Airport Gate Usage with the Critical Path Method. *Transportation Engineering Journal of ASCE*, 97(2), 187-203.

Cheng, Y. 1998a. Network-based Simulation of Aircraft at Gates in Airport Terminals. *Journal of Transportation Engineering*, 124, 188-96.

Cheng Y. 1998b. A Rule-Based Reactive Model for the Simulation of Aircraft on Airport Gates. *Knowledge-Based Systems*, 10, 225-36.

Cormen, T.H., Leiserson, C.E., Rivest, R.L., and Stein, C. 2001 *Introduction to Algorithms*. The MIT Press, Harvard.

Ding, H., Lim, A., Rodrigues, B., and Zhu, Y. 2005a. The Over-Constrained Airport Gate Assignment Problem. *Computers and Operations Research*, 32, 1867-1880.

Ding, H., Lim, A., Rodrigues, B., and Zhu, Y. 2005b. New Heuristics for the Over-Constrained Airport Gate Assignment Problem. *Journal of the Operational Research Society*, 32(7), 1867-1880.

Glover, F., and M. Laguna. 1997. *Tabu Search*. Kluwer Academic Publishers, Boston.

Gosling, G.D. 1990. Design of an Expert System for Aircraft Gate Assignment. *Transportation Research-A*, 24(1), 59-69.

Gu, Y., and Chung, C.A. 1999. Genetic Algorithm Approach to the Aircraft Gate Reassignment Problem. *Journal of Transportation Engineering*, 125(5), 384-389.

Haghani, A., and Chen M.C. 1998. Optimizing Gate Assignments at Airport Terminals. *Transportation Research A*, 32(6), 437-54.

Hassounah, M.I., and Steuart, G.N. 1993. Demand for Aircraft Gates. *Transportation Research*, 1423, 22-33.

Hu, X.B. and Chen, W.H. 2005. Genetic Algorithm Based on Receding Horizon Control for Arrival Sequencing and Scheduling. *Engineering Applications of Artificial Intelligence*, 18(5), 633-642.

Kirkpatrick, S., Gellatt, C.D., and Vecchi, M.P. 1983. Optimization by Simulated Annealing. *Science*, 220(4598), 671-80.

Lam, S.H., Cao, J., and Fan, H. 2002. Development of an Intelligent Agent for Airport Gate Assignment. *Journal of Air Transportation*, 7(2), 103-114.

Mangoubi, D., and Mathaisel, R.S. 1985. Optimizing Gate Assignments at Airport Terminals. *Transportation Science*, 19, 173-88.

Obata, T. 1979. *The Quadratic Assignment Problem: Evaluation of Exact and Heuristic Algorithms*. Technical Report TRS-7901, Rensselaer Polytechnic Institute, Troy, New York.

Ruperto, F., and Francesc, R. 2007. Reducing Total Time in the Gate Assignment Problem, Presented at *Transportation Research Board Annual Meeting 2007*, CD-Rom, Paper Number 07-1384.

Srihari, K., and Muthukrishnan, R. 1991. An Expert System Methodology for an Aircraft-Gate Assignment. *Computers and Industrial Engineering*, 21(1-4), 101-105.

Wirasinghe, S.C., and Bandara, S. 1990. Airport Gate Position Estimation for Minimum Total Costs—Approximate Closed Form Solution. *Transportation Research B*, 24B(4), 287-297.

Yan, S., and Chang, C-M. 1998. A Network Model for Gate Assignment. *Journal of Advanced Transportation*, 32(2), 176-89.

Yan, S., and Huo, C-M. 2001 Optimization of Multiple Objective Gate Assignments. *Transportation Research Part A: Policy and Practice*, 35, 413-432.

Xu, J., and Bailey, G. 2001. The Airport Gate Assignment Problem: Mathematical Model and a Tabu Search Algorithm. In *Proceedings of the 34th Hawaii International Conference on System Sciences*, Island of Maui, Hawaii, USA.

Zhu, Y., Lim, A., and Rodrigues, B. 2003. Aircraft and Gate Scheduling with Time Windows. In: *Proceedings of the 15th IEEE International Conference on Tools with Artificial Intelligence*, Sacramento, California, USA.

This page has been left blank intentionally

Chapter 8
Baggage Handling

Introduction

Baggage handling at an airport usually involves three main functions: 1) moving bags from the check-in area to the departure gates; 2) transferring bags from one gate to another; and 3) moving bags from the arrival gates to the baggage-claim area. The process starts by receiving luggage from the travelers at the outside curb or at the ticketing counters. Each received piece of luggage is given an identification tag that indicates its itinerary with a unique bar code. In large airports, the luggage pieces are placed on a transport belt toward the baggage-handling (sorting) facility. Similarly, the transfer luggage that arrives at the airport is also received and either transported directly to its outbound flights or to the airport's baggage sorting facility. This chapter presents the main issues of baggage handling at the airport: first, the management of the baggage at the airport sorting facility and, second, baggage transfer methods.

Management of the Sorting Facility

Flight Service

As shown in Figure 8.1, a typical baggage-handling facility consists of one or more quads. Each quad is a collection of piers, where baggage for each flight is accumulated before being loaded onto the flight. Each pier is usually assigned to collect the baggage of one flight at a time. However, the baggage of large flights could also be split between two or more piers. Usually, the baggage of large flights is split into two piers. One pier is allocated to the baggage that is arriving at the destination of the flights (city baggage) and another pier is allocated to the baggage that is to be transferred to other destinations for connecting passengers (transfer baggage).

Each departing flight is assigned a service period during which the baggage for this flight is accumulated at the assigned pier(s). A flight service period usually starts one to three hours before the flight's departure time and ends 15–30 minutes before its departure time. In general, this handling period is set to match the period during which travelers are expected to check in for the flight, which is usually a function of the flight size (number of passengers) and whether it is a domestic or international flight. Typically, larger and international flights require longer service periods than smaller and domestic flights. Each pier is assigned one crew

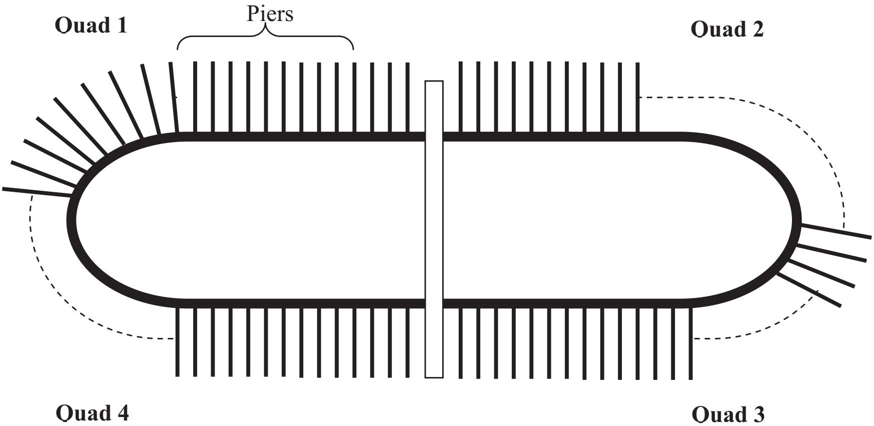


Figure 8.1 Sketch of the baggage sorting facility

member to load arriving bags on transport carts that are waiting at the pier. One crew member can simultaneously load the bags from several adjacent piers. At the end of the service period, all carts are moved to the gate to load the bags onto the aircraft of the corresponding flight.

Operational Constraints

In most cases, airports have a limited number of piers which are usually distributed among the different air carriers in proportion to their number of flights out of the airport. This means that the air carriers have to efficiently assign their flights to the available piers such that all flights receive feasible service periods. Most small and mid-size air carriers depend on the experience of their ground staff to allocate flights to the available piers. However, for a major air carrier with more than 300 daily flight departures from a major hub, this problem becomes manually intractable, and an automated tool that efficiently assigns the scheduled flights of an air carrier to the available piers in the baggage-handling facility should be used. This tool is based on a mathematical model that optimizes the utilization of available piers while satisfying the different operation constraints.

Several operation requirements are usually considered when assigning flights to available piers in the baggage-handling facility. First, the number of simultaneous in-operation piers should be minimal. As the number of simultaneous in-operation piers decreases, the manpower needed in the handling facility also decreases and the operating cost of the baggage-handling facility is consequently minimized. Second, the quantity of baggage (usage time) assigned to the different piers should be comparable. This consideration guarantees that crew members assigned to these piers receive an equivalent workload. Third, the service periods of two consecutive flights on the same pier should be

separated by a buffer period. A longer buffer period ensures that the baggage for the two flights is not mixed, and that each piece of baggage is loaded on to its corresponding flight. Fourth, if the baggage of a flight is split into multiple piers, these piers should be adjacent to each other. This arrangement reduces the transportation time required to move all carts filled with the baggage of one flight. Fifth, the piers should be assigned such that the distance between the pier and the gate of the flight is as short as possible. Again, this assignment reduces the time required to move all loaded carts from the handling facility to the flight's gate. Finally, if the air carrier is operating several daily flights to the same destination, it is preferable that these flights are assigned to the same pier. This condition allows any delayed luggage to be transported on the next flight to the correct destination without pier transfer.

Modeling Approach

Similar to the aircraft routing, crew pairing, and gate assignment problems presented in previous chapters, the pier assignment problem falls under the broader area of assigning limited resources to a number of competing prescheduled activities. Similar problems have been treated in a variety of disciplines, including machine scheduling, exhibit/lecture hall scheduling, computer processor scheduling, and so on. Although, most of these problems share similarities in their structure, the problem size and the operation constraints that govern each problem significantly differ. These constraints determine, to a great extent, the most suitable modeling approach to the problem in hand.

The principal objective of the pier assignment problem is to assign a list of prescheduled flights to available piers in the baggage-handling facility while satisfying the operation requirements described earlier. A further objective is to optimize the utilization of available piers through minimizing the number of piers simultaneously open for operation. This problem could be viewed as one version of the ASP (Cormen et al. 2001). A solution algorithm for the ASP that follows a greedy selection strategy has been proven to produce an optimal solution (Gavril 1980). This algorithm is modified to optimize the number of piers simultaneously in operation.

To illustrate the main steps of the modification of the ASP algorithm for the pier assignment problem, consider N flights that need to be assigned to P piers. Also assume that each pier is assigned to only one flight at a time, and each flight is assigned to one pier. Initially, a stack named AVAILABLE is created, and the P piers are pushed into it. In addition, a N-element array PIERS is initialized. An element PIERS[i] defines the pier number to which flight i is assigned. Finally, we define the array EVENTS of size $2N$, which stores the start and end times of the service period of the N flights. Events in this array are sorted chronologically in ascending order. In the case of a tie, the end event is placed first. In addition, a flag is associated with each element to describe whether this event represents the

start or end of a service period. Given the list of events sorted chronologically in descending order, the ASP algorithm can be written as follows;

For i =1 ***to*** 2*N*
If EVENTS[*i*].flag = start
    PIERS [*i*] = POP[AVAILABLE]
Else If EVENTS[*i*].flag = end
    PUSH[AVAILABLE, PIER[*i*]]
Return PIERS

Figure 8.2 illustrates the application of the algorithm for a problem of five flights and two piers. The algorithm scans the events in chronological order. For a start event, it allocates a pier from AVAILABLE and places it in PIERS. For an end event, it returns the pier to the AVAILABLE stack. For the example given in Figure 8.2, piers 1 and 2 (P1 and P2) are initially available in AVAILABLE. The first element in EVENTS is the start time of the service period of flight F1. Thus, F1 is assigned to P1, and this pier is removed temporarily from AVAILABLE. Then, F2 is assigned to P2, which is the next available pier in the stack. Similarly, P2 is also removed temporarily from AVAILABLE. The next element in EVENTS is the end of the service period of F1. Thus, F1 is pushed again in AVAILABLE. Flight F3 is then assigned to P1, and accordingly P1 is removed temporarily from AVAILABLE. Pier P2 is pushed again in AVAILABLE after the service period of F2 is finished. The algorithm continues until all five flights are assigned as shown in the figure.

Although the algorithm implements a greedy assignment strategy, it can be shown that the algorithm is correct and optimal. This strategy requires showing that the algorithm does not allocate more piers than necessary. In addition, each assigned pier is optimally utilized (minimum idle time). As described above, the algorithm uses a stack data structure, with a last-in-first-out queue discipline, to store all piers that are ready for assignment. Initially, the top element of this stack could be any of the available piers in the baggage-handling facility (none of them are utilized yet). While the algorithm is running, two cases could be encountered: 1) the top element in the stack is one of the previously used piers that is placed in AVAILABLE after finishing an activity; and 2) if all previously used piers are currently busy, a new unused pier is on the top of the stack. Thus, the algorithm always starts by assigning piers that are previously used. If none of these piers is available, a new pier is assigned.

The Pier Assignment Model

The algorithm described above is modified (Abdelghany et al. 2006) to consider the different operation requirements associated with the baggage-handling process. As mentioned earlier, it is preferable for piers to have comparable workloads. However, the proposed stack of available piers, AVAILABLE, uses a last-in-first-

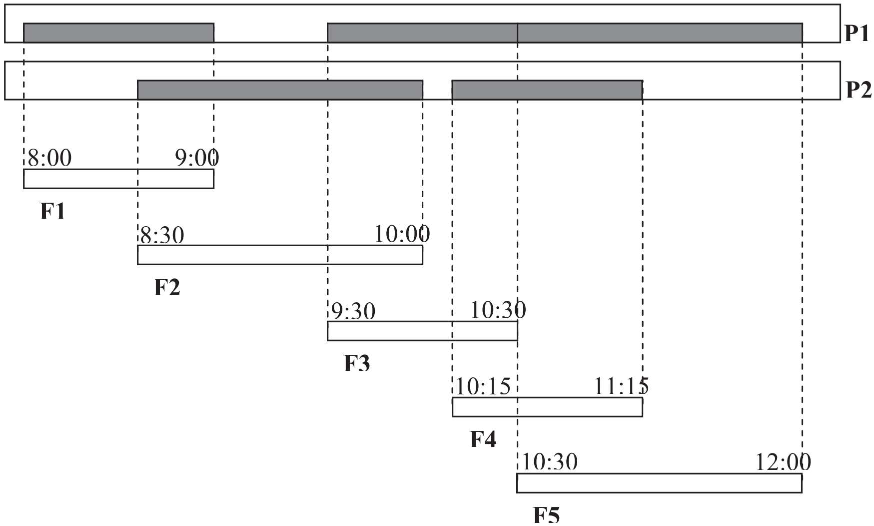


Figure 8.2 Illustrative example of the activity selection algorithm

out queue discipline. Accordingly, if several previously used piers are available, the latest released pier (the pier on top of the stack) is always selected first. This pier would likely get less idle time than all preceding piers in the stack. To avoid this happening, the top operation associated with the stack AVAILABLE is modified. Instead of simply selecting the top element in AVAILABLE, the earliest pushed pier that matches the requirements of the current flight is selected from among all previously used piers in AVAILABLE. If no previously used piers exist, a new unused pier is selected.

In addition, information on the quad of each pier is stored, enabling the assignment of flights to piers that are as close as possible to the flight's departure gate. This reduces the required time to transport all loaded carts to the flight's departure gate. Furthermore, the pier selection step is modified to consider the pier-location preference. The earliest pushed pier is selected from among the previously used piers in AVAILABLE that satisfy the location preference. If no previously used piers exist, a new pier that satisfies the location preference is selected. Then, if no pier that satisfies the location preference of the flight is found, the earliest pushed pier is selected from among the previously used piers in AVAILABLE that do not satisfy the location preference. Finally, if there is no such pier, a new pier that does not satisfy the location preference is selected.

As previously mentioned, it is recommended that flights of the same destination be assigned to the same pier so that any delayed luggage can be sent to its destination on subsequent flights without a pier change. Therefore, the information about the list of destinations served until now by each pier is stored. In the pier-selection step of the algorithm, if more than one pier is available for a flight, the pier that previously

served this flight's destination is selected. In the current implementation, no pier is kept idle while waiting to be assigned a future flight with a destination that matches the destination of flights already assigned on this pier. The effect of constraining the number of destinations allowed per pier is presented hereafter.

Finally, to ensure that a reasonable buffer exists between successive service periods on the same pier, the start (end) time of each service period in EVENTS is modified. For instance, if a buffer interval of length T is required, this interval is subtracted from (added to) the start (end) of each flight's service period. The modified start (end) times are then stored in EVENTS.

Figure 8.3 describes the pier assignment model. It starts by reading the list of outbound flights and the list of available piers. Given the start and end times of the service period of each flight, the array EVENTS is generated. Also, all piers are pushed in AVAILABLE. Elements in EVENTS are iterated. If a start of a service period is encountered, the pier selection module is activated. The pier search module finds a pier that (1) has the longest idle time interval among all previously used piers in AVAILABLE. Otherwise, a new pier is opened for operation that (2) previously served the destination of the current flight and (3) satisfies the location preference of the flight. In addition, if more than one pier is required for this flight, another adjacent pier has to be available. Once a pier(s) is found for this flight, it is stored in PIERS[i], where i is the flight index. If the next element in EVENT represents an end of a service period, the pier is pushed back to AVAILABLE. The process continues until all events are scanned, and PIERS is populated for all flights.

Clearly, infeasibility could be encountered if the number of flights to be assigned simultaneously is greater than the number of available piers in the baggage-handling facility: in other words, if AVAILABLE is empty when more flights need to be assigned. In this case, all piers are scanned to determine the pier with the earliest release time after the scheduled start of the service period of the new flight. Assigning the flight to this pier results in an overlap period (two flights simultaneously occupying the same pier). This overlap period is calculated as the difference between the start time of the new flight's service period and the release time of the pier. If this overlap is less than a predefined threshold, the service period of the current flight is shortened to eliminate the assignment unfeasibility. Otherwise, this new flight is marked as open, indicating that no pier is assigned to this flight. As a user option, the model allows this overlap interval to be distributed among all flights already assigned to this pier. The time interval to be reduced from the different flights is the length of the overlap period divided by number of flights assigned until now to this pier.

Baggage Transfer

Types of Transfer Baggage

The baggage on a flight arriving at the airport is usually classified into city baggage and transfer baggage. The city baggage ends at the destination of the flight and is

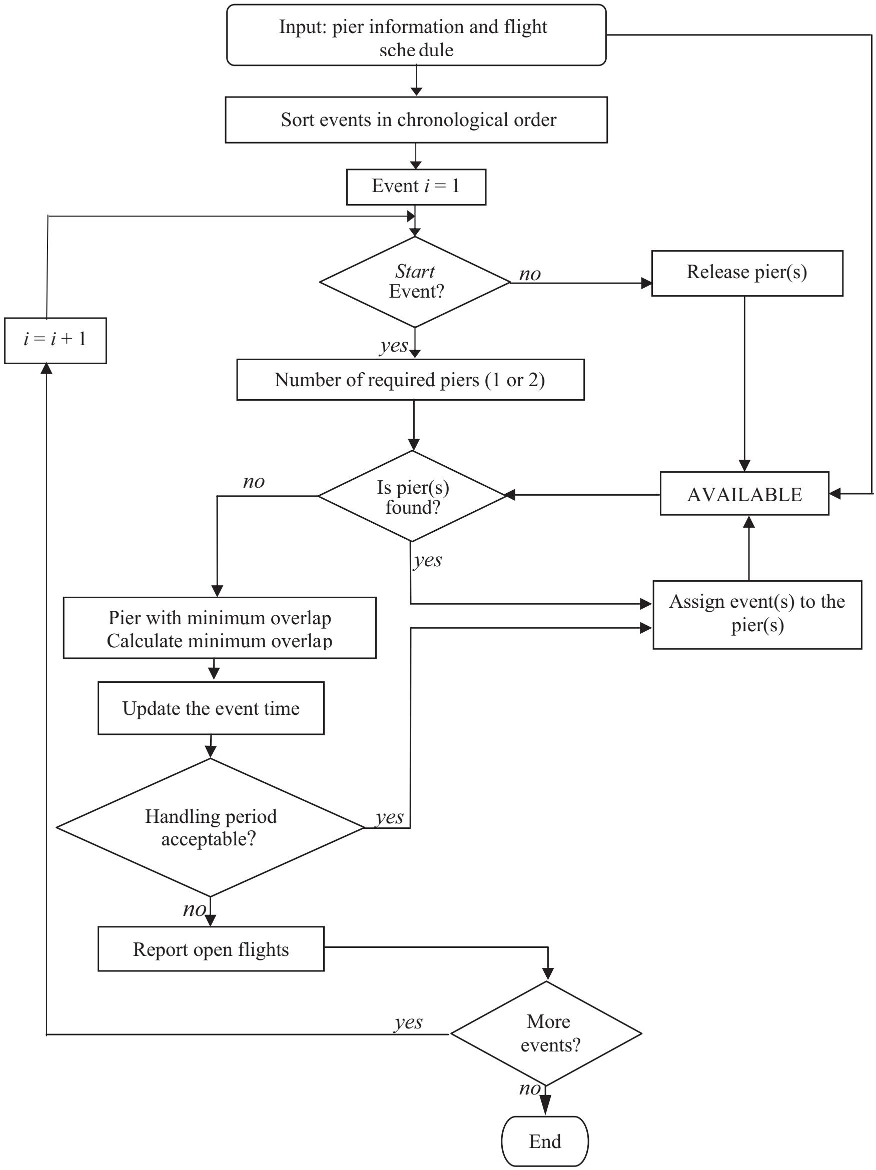


Figure 8.3 General framework for the baggage-handling model

moved to the baggage-claim area of the airport. The transfer baggage is classified further, based on the departure time of its next flight, into cold baggage and hot baggage. Cold baggage has enough time to be unloaded and moved to the sorting facility. At the sorting facility, it is accumulated with the transfer baggage arriving from other flights and the baggage starting at this airport, and then transferred to

its connecting outbound flights by transportation carts. The hot baggage, however, has no time to be moved to the sorting facility. Accordingly, it has to be moved quickly to its outbound connecting flights by transportation carts.

The transfer of the hot baggage from the inbound flight to the different outbound flights is a time-critical process that should be optimized to minimize the total transportation time and cost. The total cost of this process depends on the location of the gate of the inbound flight and the location of the gates of the outbound flights. It also depends on the departure time of the different outbound flights. Figure 8.4 shows an example of the trajectories of the transferred hot baggage of an inbound flight.

Usually, there are several inbound flights that arrive within a short time window at the airport. The hot baggage of these flights needs to be transferred instantaneously to several outbound flights that are also departing within a short time window. Accordingly, a good logistic plan needs to be set to perform the baggage transfer of all inbound flights efficiently with minimum resources.

Transfer Methods

Currently hot baggage at the airport is usually transferred using the push method or the pull method, or a combination of both. In the push method, the transportation carts wait at the gate of the inbound flight. Once the flight arrives, the hot transfer baggage is loaded on to these carts. The transportation carts are then moved to drop the baggage at their different corresponding outbound flights. The order in which the transportation carts visit the gates of the different

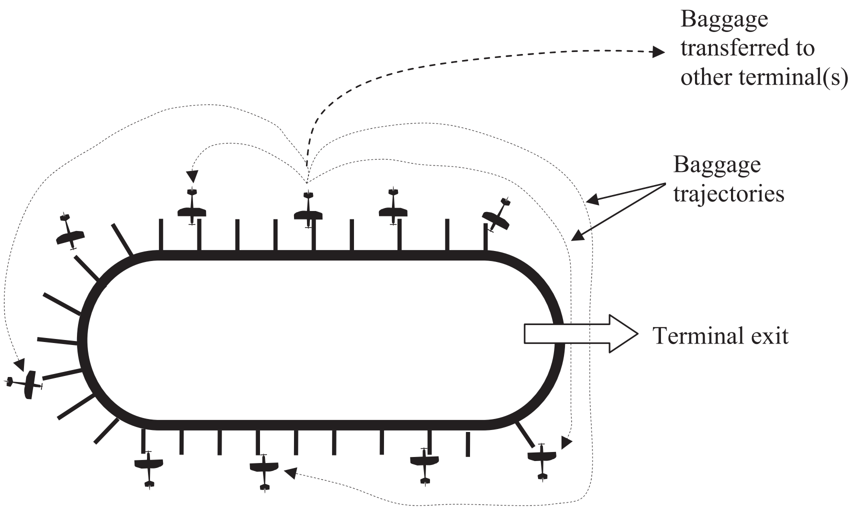


Figure 8.4 Example of trajectories of connecting baggage at the airport terminal

outbound flights depends on the location of these gates and the departure time of the different outbound flights. In the pull method, each outbound flight sends transportation carts to collect the transfer baggage from each inbound flight. Similarly, the order in which the transportation carts visit the gates of the different inbound flights depends on the location of these gates and the arrival time of the different inbound flights.

There is little attention given in the literature on how to optimize the routes of the transportation carts to efficiently minimize the total transportation cost. We only refer to the work of Green et al. (2007), who present an effort to improve the transfer baggage operation at DFW for American Airlines. A simulation model is developed to study the current practices and identify opportunities for improvement, evaluate the alternative policies and processes, and recommend changes to improve the baggage delivery performance. Using the simulation model, they test two routing scenarios of the push method against optimized routing. In the first routing, the tail-to-tail baggage is delivered in ascending order of gate number. The advantage of this routing is that it is easy to understand and easy for the drivers to follow. The main disadvantage is that departure time is not taken into account, and more baggage is at risk of missing its connection. In the second scenario, the tail-to-tail baggage is delivered in ascending order of departure time of the outbound flight. The advantage of this method is that the baggage with shorter connect time has less risk of missing its outbound flights. However, the disadvantages include the very long travel distance; the driver taking longer to return to the gate; and the fact that the plan is not easy to follow by drivers. Finally, an optimized routing that considers the location of the gates of the outbound flights as well as the flights' departure times is compared to the previous two strategies. Green et al. (2007) do not present the details of the algorithm to develop the optimal routing. While still not driver-friendly, it provides the best performance of the transfer process by minimizing not only the total delay of the baggage, but also the required resources.

Primary Contributions

There is very little literature on modeling the operation of baggage-handling facilities at airports. Early examples of the available literature include the work of Robuste (1988), who analyzed the baggage-handling operations at airports. Hsu (1993) performs a simulation study for the demand at the baggage room to forecast the room capacity. Similarly, Jaroenpuntaruk and Miller (1995) develop and use a stochastic simulation model to evaluate four alternative baggage-handling systems for American Airlines at O'Hare airport. The model provides information to decision-makers on how the baggage room operates under a given design plan and operating policy. A sensitivity analysis is performed to obtain a system-wide view of the effect of changes in baggage transfer volume.

The discussion in this chapter is based on the recent work of Abdelghany et al. (2006), who present a heuristic-based model for scheduling baggage-handling facilities in congested airports. The model assigns the baggage of departing flights to available piers in the baggage-handling facility and adopts an activity selection algorithm that is generally used to assign limited infrastructure resources to a number of prescheduled competing activities. The model considers various operational requirements associated with the baggage-handling process and is applied to a daily schedule of a major US air carrier at one of its major hubs. The results illustrate the need to consider tradeoffs between satisfying the different operation requirements and minimizing the operation cost at the facility.

References

Abdelghany, A., Abdelghany, K.F., and Narasimhan, R. 2006. Scheduling Baggage Handling Facilities in Congested Airports. *Journal of Air Transport Management*, 12, 76-81.

Cormen, T.H., Leiserson, C.E., Rivest, R.L., and Stein, C. 2001 *Introduction to Algorithms*. The MIT Press, Boston, MA.

Gavril, F. 1980. Algorithms for Minimum Coloring, Maximum Clique, Minimum Covering by Clique, and Maximum Independent Set of a Chordal Graph, *SIAM Journal of Computing* 1, 180-187.

Green, T., Scalise, A., and Tansupaswatdikul, P. 2007. Improving the Transfer Baggage Operation at DFW. *AGIFORS Operations Study Group*, Bangkok, Thailand.

Hsu, C.C., 1993. A Simulation Study of Airport Bag Room Demand to Forecast Capacity. MS thesis, University of Illinois at Chicago, IL.

Jaroenpuntaruk, J. and Miller, F. 1995. A Stochastic Computer Simulation of an Airport Baggage Facility. *Journal of Quality in Maintenance Engineering*, 1(2), 60-68.

Robuste, F. 1988. Analysis of Baggage Handling Operations at Airports. Ph.D thesis, University of California, Berkeley, CA.

Chapter 9
Flight Planning and Fuel Management

Introduction

Flight planning is the process of identifying a flying plan for each flight prior to its departure and involves two major functions. The first is to calculate the amount of fuel needed by an aircraft before flying from an origin airport to a destination airport. The second is related to determining a flight path that complies with air traffic control (ATC) requirements. The flight path and fuel estimation are determined in order to guarantee flight safety and also to minimize the cost by choosing the appropriate route, altitude, and speed, passenger and cargo load, and loading the necessary fuel on board. On international flights, air carriers are charged when flying through the airspace of other countries. Therefore, flight planning should also consider these charges (overflight costs) when deciding on the flight path and the corresponding required fuel amount.

The main input to the flight planning process is an accurate weather forecast, including wind and air temperature. The wind direction (that is, head or tail wind) affects the fuel consumption of the flight along its path. To comply with safety regulations, each aircraft has to carry more fuel than the minimum needed to fly from the origin to the destination. This requirement is to account for any unexpected circumstances such as flight diversion to another airport should the planned destination airport become unavailable.

An aircraft flying between any two airports is supervised by ATC and must follow predetermined routes known as airways. There are usually several possible paths between any two airports. Along the airways, the aircraft must maintain flight levels (altitude), depending on the route being flown and the direction of travel. The objective of commercial air carriers is to minimize flying costs through the optimum selection of route, speed, and altitude, while at the same time maintaining safety standards.

Finding an accurate optimized flight plan for each flight in the schedule requires a large number of calculations with sophisticated algorithms. Some commercial air carriers have implemented their own internal flight planning systems, while other air carriers use the services of external planners (for example, SITA, Jeppesen, and Lufthansa Systems). Though aviation regulations vary by country, most require a licensed flight dispatcher or flight operations officer to carry out flight planning for airlines.

Route and Altitude Structure

An airway can be thought of as a highway in the sky that connects between two predefined points in the airspace known as waypoints (fixes). There is a huge route

network that involves a large number of named official airways along which aircraft fly between any two airports, under the supervision of ATC. On an airway, aircraft fly at different heights to avoid collisions. Maps and charts that show airways are published by various suppliers. They are usually updated monthly, corresponding to Aeronautical Information Regulation and Control (AIRAC). Each airway starts and finishes at a waypoint and may also contain some intermediate waypoints, and airways can only cross or join at a waypoint. At these cross and joint points, an aircraft can change from one airway to another. Most waypoints are classified as compulsory reporting points, where the pilot reports the aircraft position to ATC as the aircraft passes these points.

Problem Description

As mentioned above, the objective of flight planning is to determine the aircraft's optimal flight path with the corresponding fuel load. The flight planning problem can be formulated mathematically as follows (Altus 2007):

Minimize:	Fuel cost + Time-based cost + Overflight cost + Spill cost
Subject to:	Weather
	Aircraft performance
	Allowed route and altitude structure
	Schedule and operational constraints
By varying:	Route (ground track)
	Profile (altitudes along the route)
	Speed
	Departure fuel/load

The objective function is to minimize the total cost associated with the flight operation. This cost is composed of four main components. The first component is the cost that is consumed during the flight, which depends on the aircraft performance, the aircraft weight, the route, speed, and profile. Studies in aeronautical engineering have shown that the weight of the aircraft varies dramatically during the course of the flight due to fuel consumption. Also, the fuel consumption (flow) rate varies with the altitude and temperature. In addition, the fuel required to climb varies nonlinearly with the weight of the aircraft. Because the weight of the aircraft affects its optimal route, the weight of the aircraft at any point impacts on the optimal decision at that point. The second cost component is pertinent to flight delay. During adverse weather conditions, the flight plan may cause a flight to select a longer route that results in flight delay. The overflight cost is related to the charges paid to international countries when using their airspace. This cost usually differs between countries, and flight planning should avoid selecting flight routes with high overflight cost. The spill cost is the cost encountered by the air carrier when it is forced to reduce some of the flight weight to optimize the total cost.

The air carrier might spill some of the aircraft's passengers and cargo to reduce its takeoff weight.

The main constraints of the problem include weather, aircraft performance, allowed routes, and schedule and operation constraints. The wind direction and strength, as well as temperature, are very important in determining the flight path. Previous experience has shown that the wind-optimal path of a flight can be very different from the great circle path. To accumulate weather information, the forecasts for worldwide wind and temperature are distributed by the National Weather Service (NWS) and the UK Met Office (UKMO). Some vendors and air carriers enhance these forecasts based on other sources of information, including pilot reports and different forecast models. These forecasts are generally produced every six hours, and cover the next 36 hours at intervals of six hours. In each forecast, the whole world is covered using grid points located at intervals of 75 miles or less. At each grid point, the weather (wind speed, wind direction, air temperature) is supplied at nine different heights (altitudes) ranging from about 4,500 feet up to about 55,000 feet. However, an aircraft rarely flies exactly at weather grid points or at the exact heights at which weather predictions are supplied, so some form of horizontal and vertical interpolation is generally performed to estimate the weather conditions at non-grid points. Linear interpolation is usually used to estimate the weather conditions at the different locations between grid points.

It is well known that the amount of fuel consumed depends on the aircraft type and age. Aircraft manufacturers publish information about the fuel consumption of the aircraft during different maneuvers and under different operation conditions. The flight planning process uses this information to minimize the total fuel cost.

During the flight, the aircraft should follow all requirements of air traffic controllers to avoid mid-air collisions. In that regard, the aircraft should comply with altitude requirements, separation, path, and speed. The height (altitude) at which an aircraft can fly is determined by the rules that govern each inter-waypoint portion of an airway. Also, the total aircraft weight at any point determines the highest flight level that can be used. Cruising at a higher flight level generally requires less fuel than at a lower flight level. However, additional climb fuel may be needed to get up to the higher flight level. Also, the fuel tanks have a maximum limited capacity, which depends on the aircraft type. Another constraint is related to the schedule of the air carrier that operates the flight and the operation conditions at the origin and destination airports.

Current Practice

In most cases, the flight planning process follows an ad hoc procedure that usually involves three main steps. First, for each OD pair, a set of predefined fixed routes are determined and stored (≈ 20 routes). Next, one route from the fixed set of routes between any two departure-arrival points is selected. This selection is based on finding the fastest route by assigning the approximate

weather conditions along the route. If there are adverse weather conditions, the dispatcher has to visually find which routes avoid these conditions and then select the route that is most appropriate. The last step in flight planning uses the initially selected route to find the altitude and speed schedule. This selection is usually based on wind situation, which is used to determine the best combination of altitudes and speeds to provide the best fuel and time for the flight.

The presented approach to flight planning has a number of weaknesses and is expected to provide poor solutions to the problem. First, the network structure of airways could provide millions of possible combinations for designing routes. Supplying a limited set of fixed routes is highly suboptimal. The wind patterns on any day determine the best route, and the set of fixed routes may contain a very poor set of choices on a specific day. Second, selecting the best possible route from the stored routes is also suboptimal because the route selection, altitude selection, and speed selection are tightly coupled phenomena. Therefore, any decomposition of the process into route selection and then altitude selection will produce a suboptimal solution. Third, if a route is selected based upon approximate weather conditions at different altitudes, the actual flight profile may follow different altitudes due to the payload and performance requirements of the aircraft. Therefore, the solution may not turn out to be very good if compared with the initial assumption. Fourth, the current methodology becomes less effective if there are adverse weather conditions, so that the dispatcher has to visually select the routes that avoid these conditions, which usually leads to an inefficient selection of routes.

The challenge in optimizing flight planning comes from two sources. The first is the complexity of the business process as described above. The second is the need for a very fast-running algorithm to satisfy the workload for real-time flight planning at an air carrier. It is very difficult to develop a comprehensive multi-dimensionally optimal flight plan within such a small running time. The task facing a flight planning system becomes more complicated when there is more than one possible route between the origin and destination airports because the degree of calculation required to produce an accurate flight plan is so complicated that it is not feasible to thoroughly examine every possible route. A flight planning system must have some fast way of feasibly eliminating inferior routes and leaving just a manageable number of routes that can be investigated in detail.

Solution Approach

The best solution approach to the flight planning problem would be to perform a four-dimensional (4-D) real-time trajectory optimization that considers the aircraft route, profile (altitudes), and speed. The output of this approach would be a globally optimal solution for the flight plan. The main advantage of this method is that it captures all the collective effects among the different factors that affect the

cost of the flight plan. However, the large problem size might make this approach computationally intractable, and it would be difficult to obtain an optimal solution within an acceptable time frame.

A more commonly used approach is the decomposition of the problem into route optimization, and profile and speed optimization (Altus 2007). As a first step, the methodology tries to find the lowest cost flight route (a two-dimensional (2-D) route optimization). Then, the optimal flight profile and speed is determined for the flight. The main advantage of this solution methodology is that it is more computationally tractable. Also, it is easier to enforce specific restrictions. However, the proposed solution methodology does not guarantee convergence to the global optimum. Figure 9.1 presents the framework of the flight planning process using the decomposition approach.

The 2-D route optimization usually applies the shortest path algorithm such as Dijkstra's algorithm (Ahuja et al. 1993). The objective is to find the route with the minimum total wind-corrected distance. The main challenge of this step is to define the grid, which is not trivial in the case of changes in airways and waypoints. Also during this step, information is needed about wind conditions that

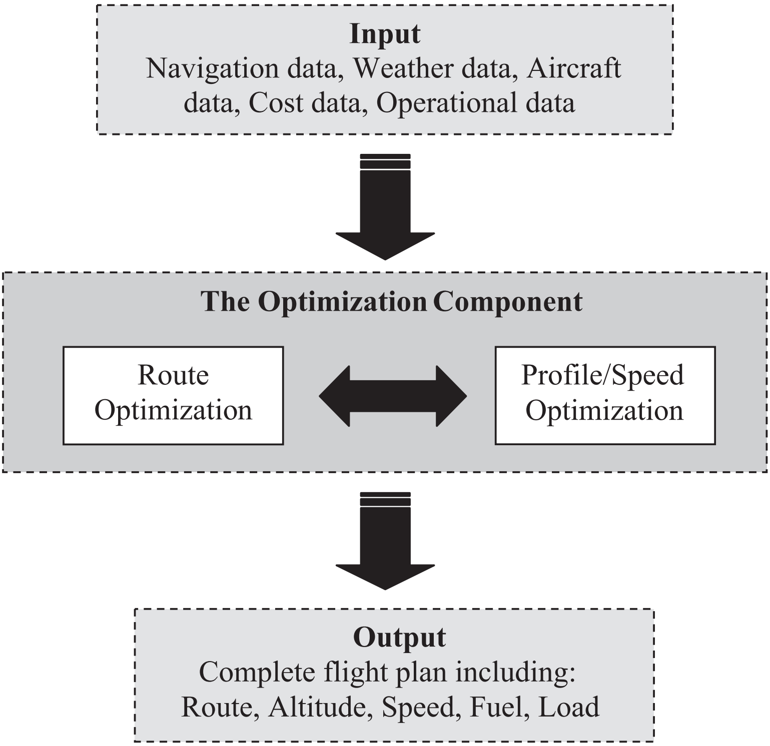


Figure 9.1 The framework of the flight planning process using the decomposition approach

are unknown because the conditions mainly depend on the altitude of the flight that is not determined yet. Several options are typically used to compensate for the missing wind and altitude information for the 2-D route optimization. The most common approach is to use a heuristic rule to select the preliminary altitude for each route (Altus 2007).

The profile optimization uses a heuristic or dynamic programming approach that considers the trade-off between the fuel load, fuel consumption, flight time, and total cost. The heuristic determines the optimal step climb location, which depends on the temperature and wind at each altitude and the aircraft weight. Figure 9.2 shows an illustrative example of the profile optimization that is given by Altus (2007).

As shown in the figure, an aircraft at 30,000 feet altitude with a total weight of 120,000 lb has two profile options. The first option is to continue to cruise at this altitude for 18.4 minutes at a fuel burn rate of 5,150 lb/hour. At the end of this maneuver, the aircraft loses 1,580 lb as fuel burn, and it can climb with less weight (118,420 lb) to an altitude of 33,000 feet. This climbing maneuver at the current weight requires 670 lb of fuel. Therefore, the total fuel consumption for the first profile is 2,250 lb. In the second profile, the aircraft starts to climb with a total weight of 120,000 lb. Since the aircraft is heavier, the climbing maneuver requires more fuel, which is estimated to be 720 lb. At the higher altitude, the aircraft has a lower fuel burning rate and less speed as given in Figure 9.3. The total fuel used in the second profile is 2,270 lb, which is higher than that of the first profile. This example illustrates the importance of considering the changing weight of the aircraft after each maneuver and the impact of the weight change on the overall cost of the profile. Commercial air carriers avoid having an aircraft change height frequently. This altitude change makes passengers uncomfortable and makes it difficult to serve meals and other services on the flight. Accordingly, airlines often set some minimum time for flight level changes that can be considered in the optimization.

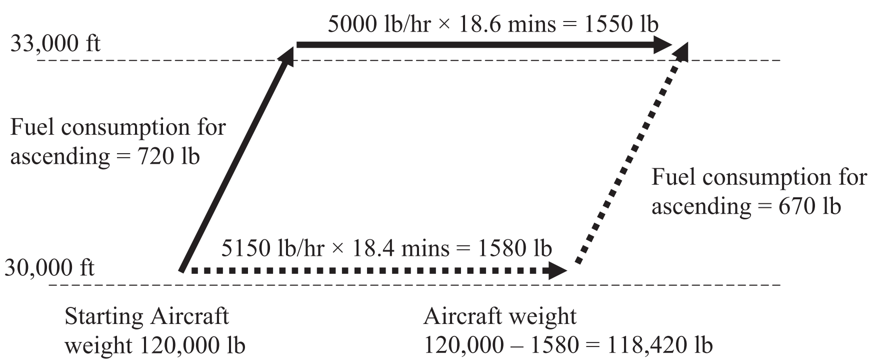


Figure 9.2 An example of profile optimization

Other Issues in Flight Planning

Reclear or Redispatch

Aircraft must carry some additional fuel to account for unforeseen circumstances, such as unexpected weathercast, or flying at a lower height than optimum due to congestion, or some last-minute passengers whose weight was not considered when the flight plan was prepared. There are several different methods that are used to estimate the additional reserve fuel, depending on the air carrier and locality. Airlines are trying to reduce the amount of reserve fuel on their flights because reserve fuel costs money. The techniques known as reclear or redispatch, or the decision point procedures that have been developed can reduce the amount of reserve fuel needed, while maintaining all required safety standards.

In these techniques, an intermediate airport to which the flight can divert if necessary due to fuel shortage is specified. Therefore, the flight plan has two destinations known as the final destination and the initial destination. The final destination airport is where the flight is really going. The initial destination airport is where the flight can divert in case that fuel shortage is anticipated. The initial destination is selected so that less reserve fuel is needed for a flight from the origin to the initial destination than for a flight from the origin to the final destination. A waypoint (usually called a reclear fix) is selected such that a decision can be made as to which destination to go to. Upon reaching this waypoint, the flight crew compares between the actual and predicted fuel burn, and checks how much reserve fuel is available. Accordingly, if there is sufficient reserve fuel, then the flight can continue to the final destination airport, otherwise the aircraft must divert to the initial destination airport. Under normal conditions, none of the reserve fuel is actually used. Accordingly, when the aircraft reaches the reclear fix, it still has all of the original reserve fuel on board, which is enough to operate the flight from the reclear fix to the final destination. The fuel savings from the reclear procedure depend on the locations of the reclear fix and the initial alternate airport.

Tankering Fuel (fuel ferrying)

As mentioned before, the route of an aircraft consists of several flights and visits different airports. When the fuel prices differ between airports, it might be economical to put in more fuel where it is cheap. However, the cost of extra trip fuel needed to carry the extra weight should be taken into consideration. A flight planning system can work out how much extra fuel can profitably be carried. Major air carriers adopt the concept of fuel ferrying to take advantage of the difference in fuel prices at different airports. Figure 9.3a–9.3c shows the fuel-loading diagram along a typical aircraft route for different fuel ferrying scenarios. A certain amount of fuel is loaded at the origin airport of each flight in the aircraft route. This fuel is then fully or partially consumed along the route of this flight. Figure 9.3a shows the fuel-loading diagram when no fuel ferrying strategy is implemented. The amount

of loaded fuel at the origin airport of a flight is fully consumed along the route of the flight. Thus, the aircraft arrives at its destination with no excess fuel. Figure 9.3b presents the fuel loading for the single-stage (flight-based) fuel ferrying strategy. Following this strategy, in addition to the amount of fuel required by a flight, an excess amount of fuel is loaded at the flight's origin airport, where the fuel price is relatively low. This excess fuel is fully consumed by the next flight in the aircraft route. Finally, Figure 9.3c describes the multi-stage (route-based) fuel ferrying strategy, where the amount of extra fuel loaded at any airport could be utilized by one or more subsequent flights along the aircraft route. Estimating the optimal amount of fuel to be carried from each station is a complex optimization problem that is described in more detail by Abdelghany et al. (2004).

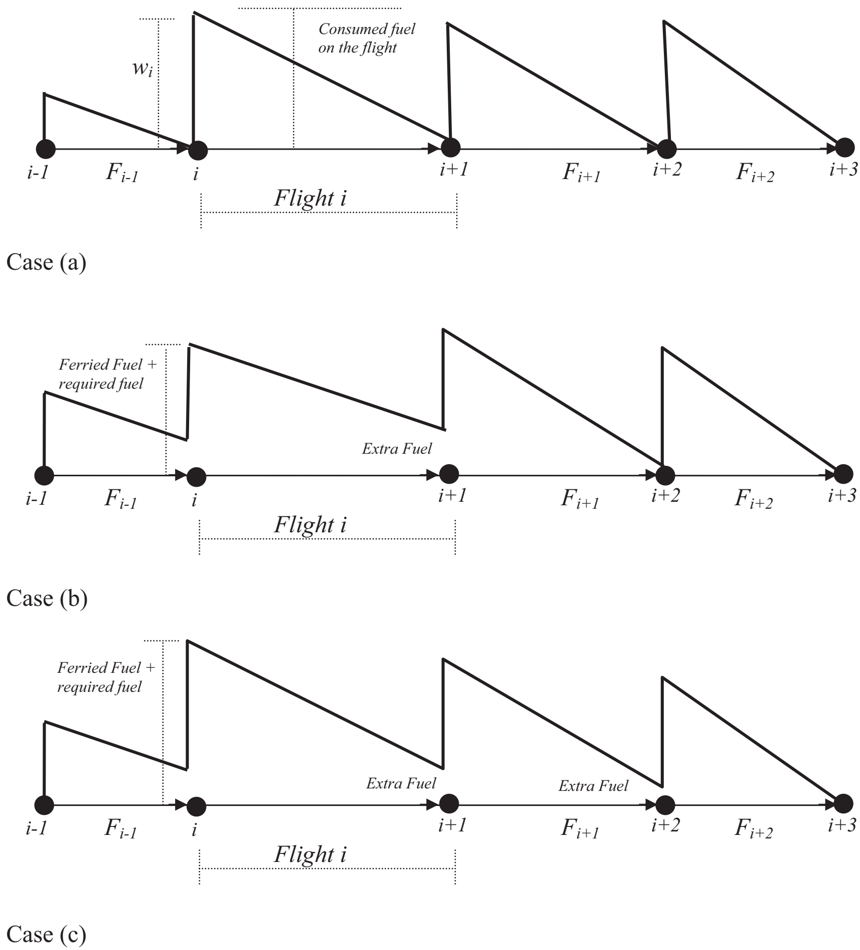


Figure 9.3 Examples of different fuel-loading patterns along the aircraft route

Primary Contributions

There is not much published literature in scholarly journals related to the flight planning problem. This is due to the commercial nature of the problem and the strong competition between vendors that supply flight planning services and software. An example of the published work is Bartholomew-Biggs et al. (2003), which present the problem of calculating aircraft fight paths to reach specified destinations while avoiding a number of obstacles and 'no fly' zones. In practical versions of the problem, the cost function used to measure the optimality of a route is likely to be non-differentiable, and experience shows that it is likely to admit multiple local minima. They discussed a number of ways of dealing with this problem, including a direct approach in which the route is determined by straight-line stages between waypoints and an indirect approach where the path is described by ordinary differential equations. Both approaches lead to formulations that involve the use of direct-search global optimization techniques.

Early efforts on modeling the airline's tankering problem are due to Darnel1 and Loflin (1977) who develop the fuel management and allocation model that determines the optimal strategy for fueling aircraft and can be used to support both short- and long-term planning. The model specifies the best fueling station and vendor for each flight, based on prices, availability, fuel burn, flight data, and cost of tankerage. Drake (1974) presented the main social, political, and economic constraints on airline fuel optimization. Barry (1981) developed a simplified model to allow Frontier Airlines to manage their fuel allocation. Linear programming techniques are utilized to develop a fuel allocation model preferable for general ease of use. Diaz (1990) formulates the airline's fuel management problem as a generalized network when there are constraints on the amount of fuel that may be purchased from a single vendor. Similarly, Stroup and Wollmer (1992) define different versions of the airlines' fuel management problem in terms of limiting the amount of fuel that may be purchased from a single airport or a single supplier. The problem formulates as a linear program. However, if there are no station or supplier constraints, it can be reduced to a pure network problem by a series of transformations on the constraints and variables. One common drawback among these models is that they ignore the trade-off between the savings associated with purchasing fuel at airports with low fuel prices and the extra fuel burn cost associated with flying the aircraft with heavier fuel weight. Teodorovic (1988) presents a strategy for the purchase of fuel on an airline network. A dynamic programming model is developed and used to determine the quantity of fuel to be bought at each airport in order to minimize the total costs of fuel purchases and satisfy the operational requirements. By comparing the results obtained after the model has been applied with the results from everyday use, the economic effects that can be achieved through an appropriate strategy of fuel purchase are evaluated. Zouein et al. (2002) present a decision support model that determines the amount of fuel to be uplifted by a plane at each station along its route over a predetermined planning horizon. The objective is to minimize the overall fuel costs. The problem

is formulated as a multiple period capacitated inventory problem and solved using linear programming. They also presented an example application that illustrates the applicability of this model to Middle East Airline's (MEA) operations. Abdelghany et al. (2004) present a model for the airlines' optimal fuel ferrying strategies. The model determines the optimal amount of fuel to be loaded at each airport for a given aircraft route. The problem is formulated using mathematical programming with an objective function that captures the trade-off between the fuel cost savings associated with loading excess fuel at airports with lower fuel prices, and extra fuel burn costs and maintenance costs associated with flying with heavier fuel weights. The benefits and limitations associated with using fuel ferrying as a cost-cutting strategy are examined.

References

Abdelghany, K.F., Abdelghany, A., and Raina S. 2004. A Model for the Airlines' Optimal Fuel Ferrying Strategies. *Journal of Air Transport Management*, 11, 199-206.

Ahuja, R.K., Magnati, T.L., and Orlin, J.B. 1993. Network Flows: Theory, Algorithms and Applications, Printice-Hall Inc., Upper Saddle River, New Jersey. Altus, S. 2007. *Flight Plannin—The Forgotten Field in Airline Operations*. Presented at AGIFORS Airline Operations, Denver, CO.

Bartholomew-Biggs, M.C., Parkhurst, S.C., and Wilson, S.P. 2003. Global Optimization Approaches to an Aircraft Routing Problem. *European Journal of Operational Research*, 146, 417-431.

Barry, N. 1981. A Simplified Alternative to Current Airline Fuel Allocation Models. *Interfaces*, 11(1), 1-9.

Darnell, D.and Loflin, C. 1977. National Airlines Fuel Management and Allocation Modeling. *Interfaces*, 7, 1-16.

Diaz, A. 1990. A Network Approach to Airline Fuel Allocation Problems. *Annuals of the Society of Logistics Engineering*, 2, 39-54.

Drake, J.W. 1974. Social, Political and Economic Constraints on Airline Fuel Optimization. *Transportation Research*, 8, 443-449.

Stroup, J. and Wollmer, R. 1992. A Fuel Management Model for the Airline Industry. *Operations Research*, 40, 229-237.

Teodorovic, D. 1988. Strategy for the Purchase of Fuel on an Airline Network. *Transportation Planning and Technology*, 12(1), 39-44.

Zouein, P., Abillama, W., and Tohme, E. 2002. A Multiple Period Capacitated Inventory Model for Airline Fuel Management: A Case Study. *Journal of Operational Research Society*, 53, 379-386.

SECTION III
Revenue Management

This page has been left blank intentionally

Chapter 10
Introduction to Revenue Management

Introduction

The substantial deregulation of air transportation in the US markets in the late 1970s provided air carriers control over most essential marketing variables such as prices, market entry and exit, capacity, and pricing. Consequently, the 1980s period witnessed the introduction of several small scheduled and charter air carriers in the different markets. These air carriers operated under a low-cost structure and were able to offer low-priced tickets in many markets. The existence of these Low-Cost Carriers (LLC) was causing revenue dilution to major pre-deregulation carriers. Major air carriers in the US. investigated possible ways to respond to this new competition. Initial thoughts were given to possible ways to cut costs to be competitive with the LCCs. However, because the load factor is low on most of the flights of the major air carriers, thoughts were diverted as to how to fill as many seats as possible on these empty flights. In an effort led by American Airlines (AA), it was decided that a limited number of seats could be offered with discounted fares to attract price-sensitive travelers. To prevent passengers willing to pay high fares from buying these discounted fares, a 21-days advance purchase restriction was added to the discounted fares. Figure 10.1 depicts a sketch of possible classification of seats into full-fare seats and discounted seats.

When AA began the discounted fare campaign, it started with a basic allocation of about 30 percent of the seats on each flight to be sold as discounted seats. However, it soon recognized that flights behave differently, and the passengers interested in discounted fares differ from one flight to another. The number of travelers interested in discounted fare varies by market type, day of week, time of day, and so on. It was soon realized that a significant effort was needed to understand the behavior of travel demand and manage the air carrier revenue at a more detailed level, which was the beginning of the techniques of air carriers' yield management, which recently has become known as the air carriers' revenue management. The goal of Revenue Management (RM) is to maximize air carrier's profit by controlling pricing and the availability of seats. It is also defined as selling the right seat to the right customer at the right price, at the right time.

Classification of Travelers

RM relies on the fact that travelers have different characteristics and needs. For example, in a broad classification air travelers are usually categorized, based on the

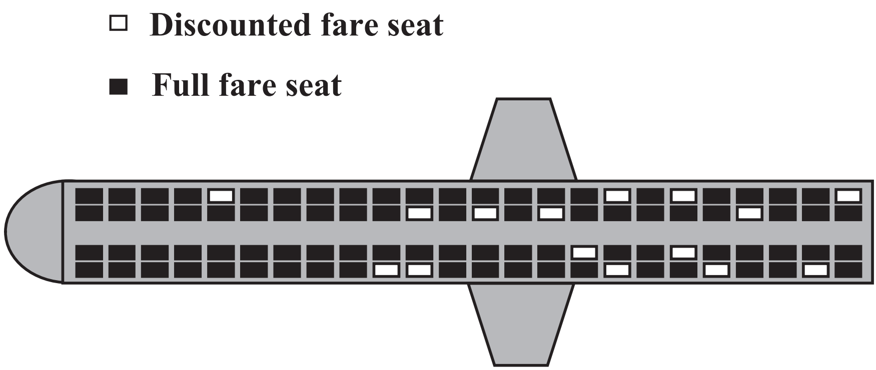


Figure 10.1 A sketch of the seats on a flight classified into full-fare seats and discounted seats

purpose of their trip, into leisure travelers and business travelers. Leisure travelers typically travel for recreational purposes or to visit families and friends. They are usually sensitive to the airfare price (elastic demand). They usually tend to book their tickets far in advance of departure. In most cases, their schedule is flexible, and usually their trip extends over the weekend. Business travelers travel to attend business meetings, conventions, and so on. In contrast to leisure travelers, the schedule of business travelers is not flexible, and they usually spend weekdays at the destinations. In many cases, they know about their business meetings at short notice, and accordingly they do not book their tickets far in advance. Business travelers are usually less sensitive to airfare price (inelastic demand). Their also prefer flexible tickets that allow last-minute itinerary change or cancellation with no penalties.

Figure 10.2 depicts the expected booking pattern for a hypothetical flight. As shown in the figure, few booking requests start as early as a few months before the flight departure date. These bookings are expected to be for leisure travelers who tend to book their tickets far in advance. The number of bookings from leisure travelers increases, as time gets closer to the flight departure date. The peak of leisure bookings is roughly four to eight weeks before departure. The number of leisure travelers decreases significantly a few days before departure. When the flight departure date gets closer, another stream of bookings materializes. These bookings are expected to be from business travelers, who usually reach their peak within two to four weeks before departure. While there is sometimes overlap between the bookings from business travelers and leisure travelers, it is important for an air carrier to differentiate the different booking requests to identify business travelers that are usually willing to pay a higher fare. In many cases, air carriers identify leisure traveler booking requests by the Saturday-stay rule. Passengers that require an itinerary that involves a Saturday stay are usually leisure travelers, and air carriers typically set lower fares on these itineraries, when seats are available.

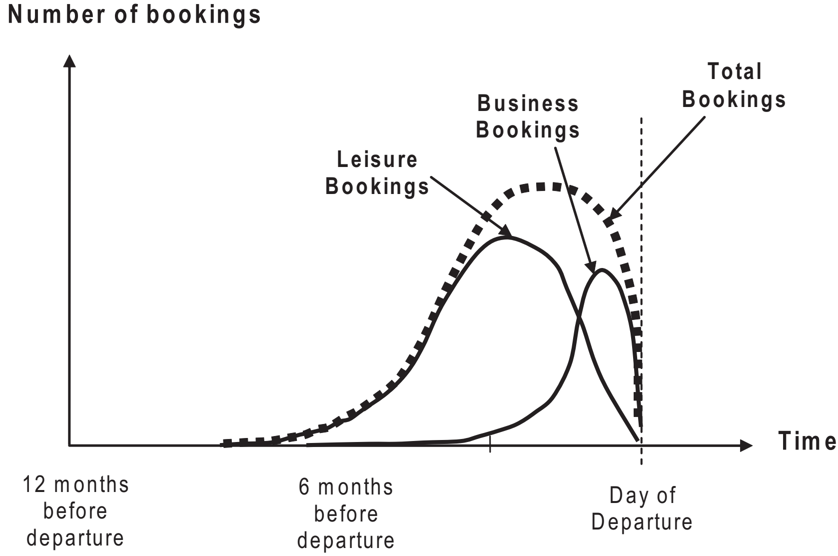


Figure 10.2 Booking pattern over time for business travelers and leisure travelers

Willingness to Pay

One major phenomenon in marketing is the fact that customers have different willingness-to-pay values for the same product. This phenomenon is represented by the demand-price curve. The demand-price curve gives the number of customers that are willing to buy the product or the service at a given price. Figure 10.3 presents a hypothetical linear-shaped demand-price curve for flight seats. As shown in the figure, as the ticket price decreases, more travelers are willing to buy seats on this flight. According to the demand-price curve shown in Figure 10.3, if the air carrier offers a \$400 fare, about 50 passengers are willing to buy seats on this flight. The total revenue at this scenario is \$20,000 (50 × \$400). If the flight capacity is 100 seats, the load factor on this flight is about 50 percent, where there are about 50 empty seats. To attract price-sensitive demand, the air carrier might offer another fare at a lower price for this flight. For instance, in Figure 10.4, it is proposed that an additional fare of \$200 is offered where there are about 75 travelers willing to buy this fare. We assume that the 75 travelers who are willing to buy the \$200 fare are independent from the 50 travelers who are willing to buy the \$400 fare. None of these 50 travelers diverts to buy the lower-priced fare. In this case, there are 125 travelers (50 + 75) that are willing to buy seats on the flight.

Since the flight capacity is only 100 seats, the air carrier has to make sure that enough seats are reserved for travelers that are willing to pay the higher fare, which is 50 seats. The remaining 50 seats are sold to 50 travelers from among the

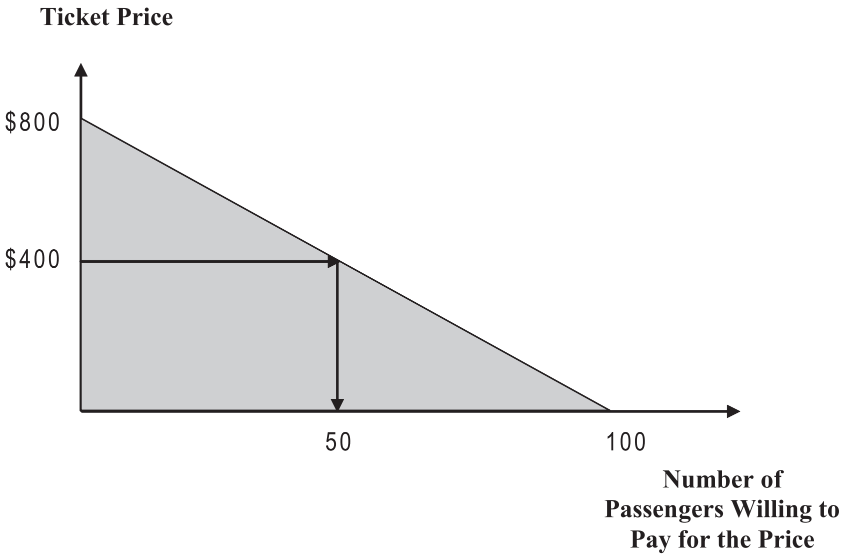


Figure 10.3 An example of a demand-price curve for flight seats (one fare)

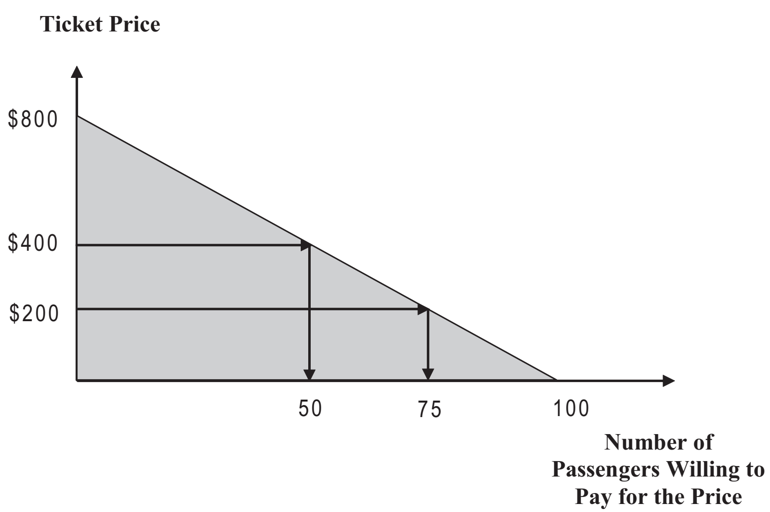


Figure 10.4 An example of a demand-price curve for flight seats (two fares)

75 travelers that can only afford the lower fare. This strategy makes the air carrier reject booking requests for 25 passengers from these passengers. The total revenue of the air carrier in this scenario is \$30,000 (50 × \$400 + 50 × \$200). In order to maximize its revenue, the air carriers must find the optimal fare mix to offer to customers. Also, it should make sure that the passengers willing to buy higher-priced fares are not diverted to buy the available lower-priced fare. This can be

achieved by adding restrictions to the lower-priced fare such as advance purchase or Saturday-stay. Finally, at all times during the booking horizon of the flight, an adequate number of seats should be reserved and protected for high-revenue passengers willing to buy the high-priced fare.

Demand Forecasting and Seat Inventory Control

To explain the main idea behind RM, consider the decision scenario given in Figure 10.5, which represents selling one seat on a flight operated by the air carrier. In this scenario, there is a booking request for a discounted fare ($T). If the air carrier accepted this request, the revenue from selling this seat would equal $T. The air carrier might decide to reject this current booking request, hoping that in the future a last-minute business traveler might be interested in buying this seat at a higher rate ($Y). If the seat is sold to the business traveler, the revenue out of selling the seat would equal ($Y > $T). However, there is a risk that no business traveler is interested in buying the seat, and the flight will depart with the seat empty. In this case, there would be no revenue from this seat.

The question is if the air carrier would decide on either selling the seat as a discounted-fare seat or take the risk and protect the seat for the higher revenue business traveler. The decision of the air carrier depends on two main factors. The first factor is the difference between the two fares $Y and $T. If there is no large difference between the values of these two fares, the air carrier should not take the risk of rejecting the booking request for the discounted fare $T. The second factor is related to the chance that a business traveler shows up late and is interested in buying the seat as a full-fare seat ($Y). For example, in the extreme case, if the air carrier is 100 percent sure that a business traveler would be interested in the full-fare seat, the air carrier should reject the booking request of the leisure traveler and protect the seat for this business traveler. However, what if the air carrier is

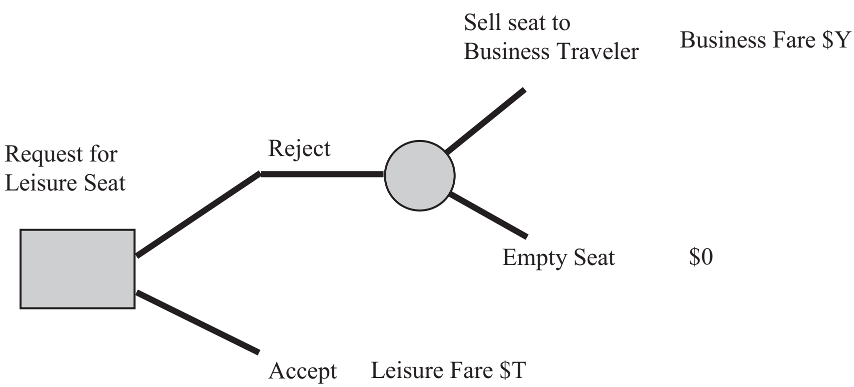


Figure 10.5 Possible decisions of selling a seat on a flight

only sure by a 50 percent or 60 percent probability that the business traveler would show up for the seat? Also, how would the air carrier estimate this probability? Predicting whether there is a business traveler interested in buying the full-fare seat is a major step in the RM process. This is referred to as demand forecasting. Demand forecasting allows the air carrier to predict the demand for each fare. Consider the case where a discounted fare \$T is sold for \$300 and the full fare Y is sold for \$500. If, for example, the air carrier is only sure by 60 percent that the business traveler is interested in buying the seat at full fare (\$Y), the air carrier can calculate the expected revenue if the seat is to be reserved for the business traveler. Two main possibilities might occur. There is a 60 percent chance that the business traveler will buy the full-fare seat and a 40 percent chance that they will not buy this seat. The expected revenue can be calculated as follows:

Expected Revenue = 0.6 × \$500 + 0.4 × \$0 = \$240

This expected revenue is less than the value of the discounted fare, which is equal to \$300 and the leisure booking request is to be accepted. Generally, the discounted fare booking request is accepted if the discounted fare is higher than the expected revenue from selling the seat later as a full-fare seat.

It can be concluded that there are two major modules for RM: demand forecasting and seat inventory control. Demand forecasting predicts the number of passengers by type on each flight itinerary. Seat inventory control allocates the seats of different flights to different groups of travelers to maximize the overall revenue.

The main objective of the RM system is to guarantee seat availability for the high-revenue passengers. This objective is achieved by protecting an adequate number of seats for those travelers and preventing the sale of these seats for cheaper fare classes. The main question is how many seats are to be reserved for those passengers. Two possible scenarios are possible:

Scenario I: What happens if too many seats are reserved for late-booking high-revenue passengers?

In this scenario, fewer seats are allowed for early-booking low-revenue passengers, and many booking requests from these travelers are rejected. If not enough late-booking high-revenue demand travelers show up, the flight departs with many empty seats. This involves revenue loss because these empty seats could have been filled with the early-booking low-revenue passengers that were rejected.

Scenario II: What happens if too few seats are reserved for late-booking high-revenue passengers?

In this scenario, many seats are allowed for early-booking low-revenue passengers. Cheap seats are widely available on the flight. Many early-booking low-revenue booking requests are accepted on the flight, and the flight is filled quickly with low-

revenue passengers. The flight load factor increases. However, selling too many seats at low fares can cause a reduction in the per-passenger revenues (yields). Also, passengers willing to pay higher fares might divert to more readily available low fares, which leads to lower revenues. Finally, since the flight is filled quickly, no seats would be available for last-minute unforeseen travelers who are willing to pay higher fares, and these high-revenue passengers would seek seats on other air carriers.

Seat Protection and Sequential Nesting

The seats on each flight are classified into several booking classes (fare classes) that are usually designated by alphabetic letters such as F (usually for First class, if available), J, U (for business class, if available), and Y, M, B, H, Q, V, W, S, T (for coach class). The seats that are within the same booking class are usually sold at the same price. Each booking class could be classified further to create additional booking classes. Booking classes are differentiated from each other by fare restrictions and rules that include:

- Saturday stay;
- advance purchase (7 days, 14 days, 21 days, and so on);
- OD of the trip;
- duration at the destination;
- refund and cancellation policy;
- one way or round trip;
- point of sale.

To define the protection level for a booking class, consider the example of a flight operating from point A to point B. As shown in Figure 10.6, this flight has 96 seats. These seats are to be distributed into two booking classes between business travelers and leisure travelers, traveling from A to B. The first booking class is designated by Y and is reserved for high-revenue business travelers (full fare). The second booking class is designated by T and is allocated for leisure travelers (discounted fare). We assume that any booking request that involves a Saturday-stay at the destination is a leisure traveler. Historical data for this flight indicates that there are about 32 booking requests for business travelers. Accordingly, a reasonable protection level for the booking class Y is to reserve 32 seats for booking class Y and the remaining 64 seats (96–32) are to be allocated for booking class T.

According to this classification, a maximum of 64 seats can be sold for booking requests that involve Saturday-stay (discounted seats). However, what is the maximum number of seats that can be sold as full-fare seats? What happens if the 32 seats allocated for booking class Y are sold and another seat booking for this booking class is requested? Revenue-maximization implies that any available seat on the lower booking class T should be sold to this additional booking request.

Therefore, theoretically, all seats available for sale in the T class should also be available for sale in the Y class. The maximum number of seats available for sale for booking class Y is 96, as shown in Figure 10.6.

Making seats that are available for sale in one booking class also available for sale in higher booking classes is known as serial or sequential nesting. In sequential nesting systems, seats that are available for sale to a particular booking class are also available to bookings in any higher-fare booking class, but not the reverse. Figure 10.7 presents another example of sequential nesting for four different fare classes on a flight. These classes, designated as Y, M, S, and T, are ordered such that Y represents the highest fare class and T represents the lowest booking class. Initially, the number of seats allocated to booking classes Y, M, S, and T are 24, 24, 28, and 20, respectively. According to the sequential nesting, the maximum number of seats available for sale for classes Y, M, S, and T are 96, 72, 48, and 20, respectively. Figure 10.8 gives another graphical representation on the sequential nesting of seats on the flight given in this example.

According to the discussion given above, we establish the following definitions:

- Protected Seats—are restricted to bookings in one or more booking (fare) classes. For instance, in the example given in Figure 10.8, we have 24 seats protected for booking class M, which are restricted to bookings in fare class M and Y only.
- Protection Level—is the total number of protected seats for a booking class. In the previous example, we have 24 seats protected for booking class Y.

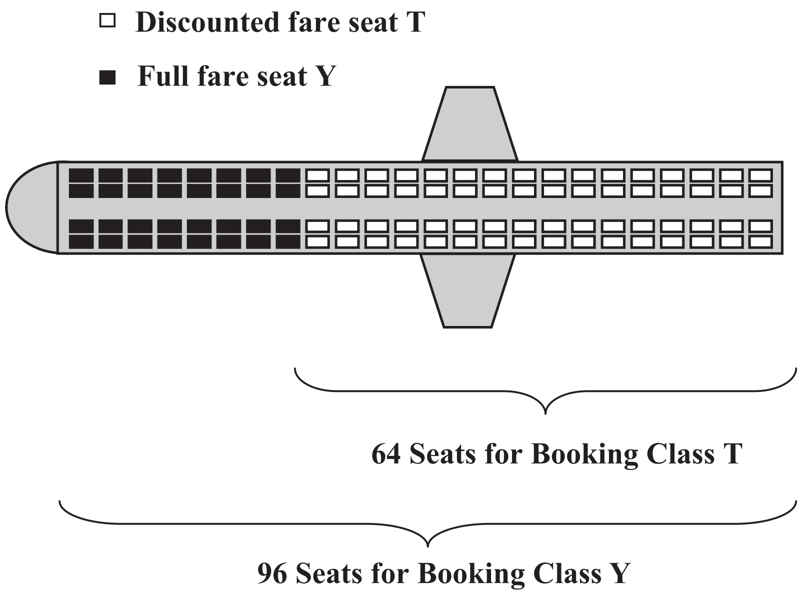


Figure 10.6 Seat availability for two booking classes

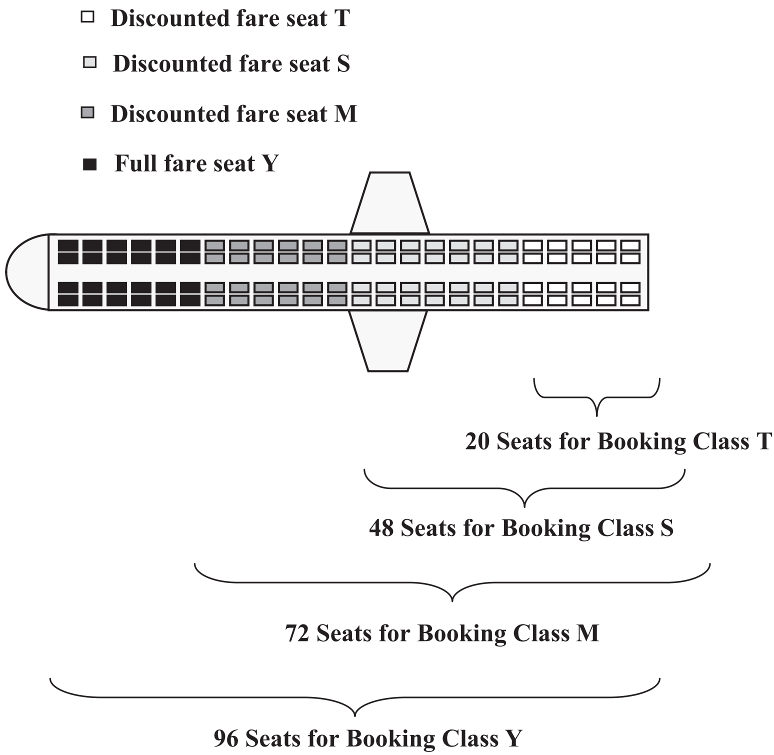


Figure 10.7 Representation of sequential nesting and seat availability for each booking class

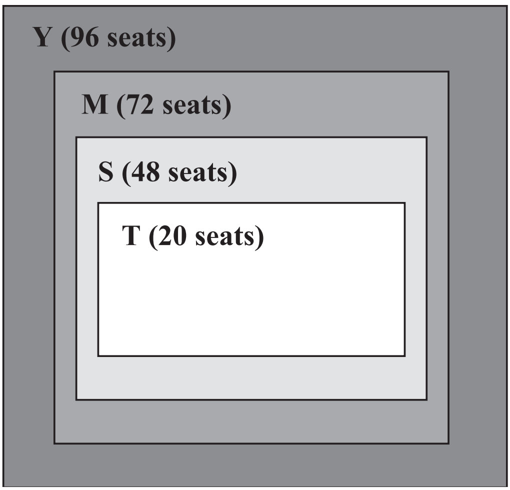


Figure 10.8 Another graphical representation of sequential nesting

- Authorization Level (AU)—is the maximum number of seats available for sale in a booking class. In the previous example, the authorization level for booking class Y is 96 seats.

Other Forms of Seat Nesting

Sequential nesting is not the only way to allocate seats of a flight to the different booking classes. For instance, in the previous example, the air carrier might decide to not to nest the seats available for fare class T within fare class S. Accordingly, the authorization level of the booking class S is only 28 seats instead of 48, as shown in Figure 10.9. In this case, the fare classes S and T are sold in parallel (parallel nesting). The parallel nesting of the S and T fare classes guarantees that there are 20 seats always available for fare class T, and those seats cannot be sold within the next upper fare class S. An air carrier might want to apply a parallel nesting structure to maintain the competition advantage in some markets. Parallel nesting allows air carriers to maintain and control a predefined number of seats to be sold for a particular group of travelers coming from special sources such as code share partners, advertising and TV promotions, corporate agreements, and so on.

Leg-based versus Network-based Revenue Management

In Chapter 1, we distinguish two different flight network structures that are adopted by air carriers, known as point-to-point and hub-and-spoke. Figures 10.10 and 10.11 present examples of the point-to-point and hub-and-spoke network structures, respectively. The point-to-point network structure is when an air carrier

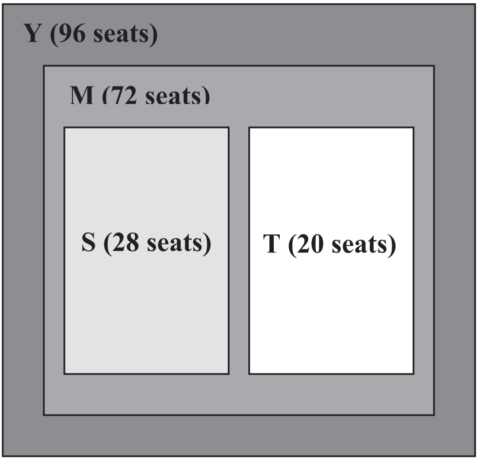


Figure 10.9 An example of a mix of parallel nesting and sequential nesting

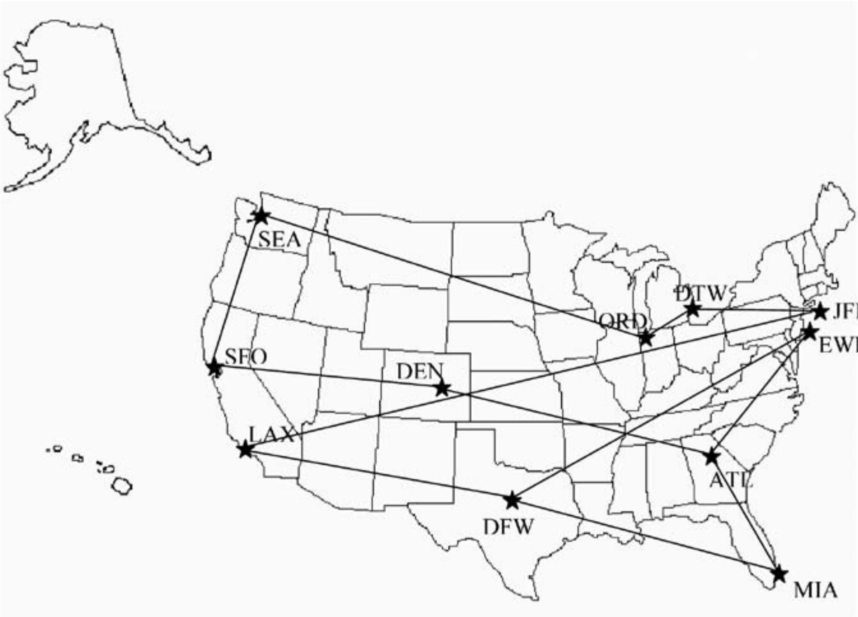


Figure 10.10 An example of a point-to-point network structure

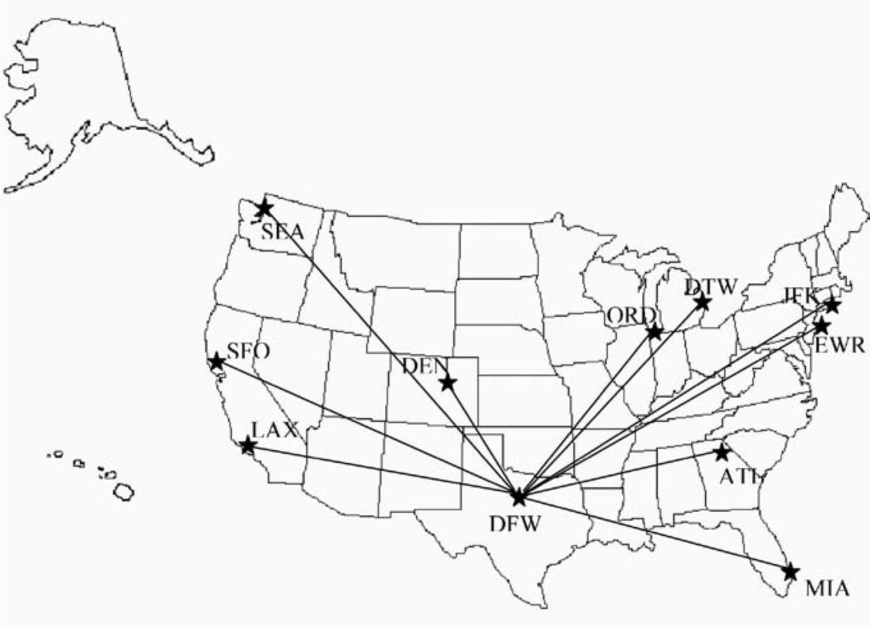


Figure 11.11 An example of a hub-and-spoke network structure

focuses on severing local traffic between city-pairs, and less attention is given to the connecting traffic to the beyond destinations. For any flight in the point-to-point network structure, the objective of the RM system is to allocate the seats of the flight among different groups of travelers (for example, business and leisure travelers) that travel between the origin and destination of the flight. On the other hand, the hub-and-spoke network structure is where every flight serves both local travelers between the origin and the destination of the flight, and connecting passengers to and from other destinations. Therefore, passengers on the flight are composed of a mix of traveler groups that differ in their trip purpose (business or leisure) and the origin and destination of the trip. For example, for the air carrier network structure shown in Figure 10.11, the flight DFW-SFO could be serving the following group of passengers:

DFW-SFO	Business
DFW-SFO	Leisure
ORD-DFW-SFO	Business
ORD-FDW-SFO	Leisure
DTW-DFW-SFO	Business
DTW-DFW-SFO	Leisure
JFK-DFW-SFO	Business
JFK-DFW-SFO	Leisure
EWR-DFW-SFO	Business
EWR-DFW-SFO	Leisure
ATL-DFW-SFO	Business
ATL-DFW-SFO	Leisure
MIA-DFW-SFO	Business
MIA-DFW-SFO	Leisure

For the hub-and-spoke network structure, the RM system is more challenging as it has to find the optimal mix of all these different groups to maximize the overall revenue of the schedule. Chapters 13 and 14 describe in detail the RM systems for airlines that adopt point-to-point and hub-and-spoke network structure, respectively.

Why Revenue Management is Applicable in the Airline Industry

Researchers have identified several characteristics for a system such that RM techniques can be applied. The first is that the product and service to be sold is perishable. The product has a time horizon in which it is available for sale. By the end of this horizon, the product cannot be sold and it has no value. This is true for a flight seat, where the seat is available for sale during the booking horizon of the flight (9–12 months). If the seat is not sold, and the flight departs, the seat has no value, and the air carrier lose its potential revenue. Second, RM is applicable

when future demand is uncertain. When air carriers protect seats for business travelers, it is not guaranteed that all protected seats will be sold. Therefore, the air carrier has to trade-off between accepting a booking request for a low fare class and rejecting this request to protect the seat for late-booking business travelers. RM is also applicable when customers have different needs and characteristics. As mentioned above, travelers on one flight might have different characteristics including trip purpose, willingness to pay, origin and destination (for hub-and-spoke air carriers), and so on. These differences allow air carriers to differentiate the product in terms of the product price. Another important characteristic for the applicability of RM is that the fixed cost of production of the product is very high relative to the marginal cost of serving a new customer. For the air carrier industry, the cost of adding a passenger to a flight does not cost the air carrier more than the meal that the passenger will consume. This cost is significantly lower compared to adding (producing) new seats to a flight that would entail a large investment in the upgrade of the aircraft size. Finally, RM is applicable when there is a fixed inventory (capacity) for sale. This is true for air carriers where each flight has a fixed seat capacity. It should be noted that other industries in which RM is applicable include hotels, car rentals, radio and television broadcasting, theaters, sport events, and newspapers and magazine advertisements.

In the next few chapters, the airline's RM is discussed in more detail. First, the main common methodologies and challenges for demand forecasting for RM are presented. Second, the overbooking problem is presented, which addresses how the actual capacity of the flight is determined. Next, the seat inventory control problem is presented for leg-based RM systems and network-based RM systems.

This page has been left blank intentionally

Chapter 11
Demand Forecasting for Revenue Management

Introduction

As explained in Chapter 10, the main idea behind Revenue Management (RM) is to guarantee seat availability for high-revenue passengers on each flight to maximize the overall revenue of the schedule. RM requires the existence of a rigorous demand forecasting mechanism to predict all demand streams on each flight in the air carrier schedule. To define the problem of demand forecasting in RM, consider the hypothetical air carrier network that is given in Figure 11.1. This air carrier has two major hubs at DEN and ORD, where there is one international flight out of ORD to London-Heathrow airport (LHR) and one international flight out of DEN to Frankfurt (FRA). According to this network, a seat on the flight DEN-ORD could be sold to a passenger in any of the following itineraries. The price of the fare on each itinerary is given in brackets.

Assume the hypothetical case that there is one seat left on the flight DEN-ORD; to which itinerary this last seat should be reserved? The answer depends on our demand forecasting for passengers on each itinerary in the network and the revenue of each passenger. For example, if demand forecasting indicates that there is a passenger interested in the itinerary LAX-DEN-ORD-LHR, the seat on the flight DEN-ORD might be reserved for this itinerary as it contributes a high revenue of $1,300.

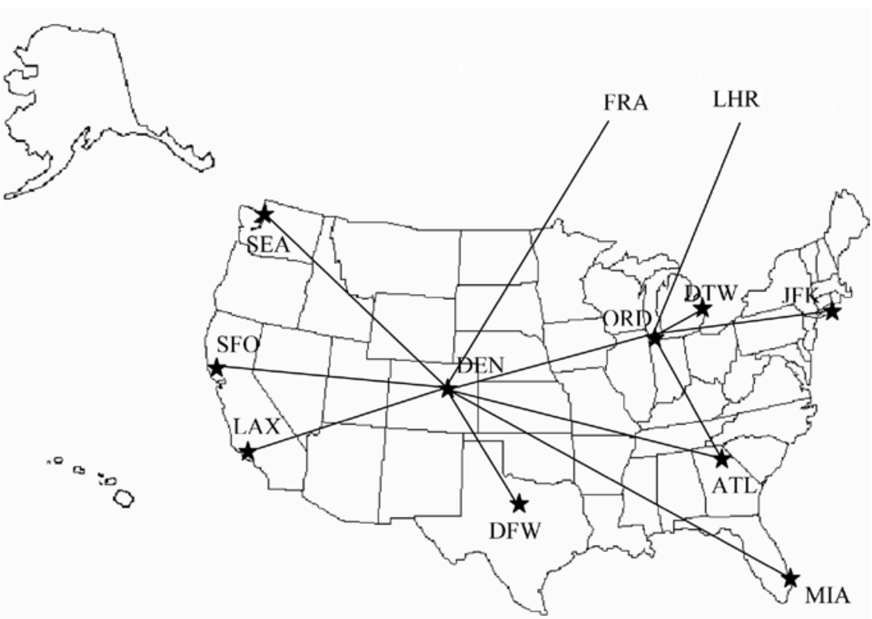
Assume also that there is one seat left on the flight LAX-DEN, and it is predicted that there is one passenger interested in the itinerary LAX-DEN-FRA at $1,500. In this case, the last seat on the flight LAX-DEN might be reserved for this passenger because it contributes $1,500 to the air carrier's revenue. By so doing, the expected passenger on the itinerary LAX-DEN-ORD-LHR is not accommodated by the air carrier, because the last seat on the flight LAX-DEN is reserved to the LAX-DEN-FRA passenger. Accordingly, the last seat on the flight DEN-ORD will be unoccupied. This seat might be reserved for a passenger on the itinerary SFO-DEN-ORD-LHR, if available, which contributes $1,250 to the air carrier total revenue. The total revenue for this seat allocation pattern is $2,750 ($1,500 + $1,250).

However, if the demand forecasting indicates that there is no passenger interested in the itinerary LAX-DEN-FRA, the last seat on the LAX-DEN flight will remain unoccupied. Accordingly, the last seat on the flight DEN-ORD should be allocated to the passenger interested in the itinerary LAX-DEN-ORD-LHR. This seat allocation pattern generates a revenue of $1,300. This example shows the significance rule of demand forecasting for air carrier RM. A rigorous demand forecasting enables the correct seat allocation and contributes to air carrier's total revenue.

Problem Definition

Consider a flight *f* in the air carrier schedule that is scheduled to depart on day *t* in the future. It is required to estimate the number of travelers Y_i^c that are interested in fare class *c* on each itinerary *i* that includes flight *f*. The different fare classes on each itinerary reflect the differences in the passengers' needs that are related to the day of purchase, Saturday-stay, change and cancellation policies, and so on. In the air carrier network example shown in Figure 11.1, the passengers of about 20 different itineraries are competing for the seats of the DEN-ORD flight. If the passengers of each itinerary are classified into four fare classes, then about 80 passenger classes will compete for the seats of the flight.

To clarify the problem size for the passenger demand forecasting, consider an air carrier that operates 1000 flights on a daily basis. Therefore, over a year period in the future, about 365,000 flights will be open for bookings and will need to be managed by the air carrier. If it is assumed that each flight is serving



SEA-DEN-ORD-ATL ($400)
SFO-DEN-ORD-ATL ($400)
LAX-DEN-ORD-ATL ($400)
DEN-ORD-ATL ($300)
SEA-DEN-ORD-DTW ($400)
SFO-DEN-ORD-DTW ($400)
LAX-DEN-ORD-DTW ($400)
DEN-ORD-DTW ($300)
SEA-DEN-ORD-LHR ($1,200)
SFO-DEN-ORD-LHR ($1,250)
LAX-DEN-ORD-LHR ($1,300)
DEN-ORD-LHR ($1,000)
SEA-DEN-ORD-JFK ($400)
SFO-DEN-ORD-JFK ($400)
LAX-DEN-ORD-JFK ($400)
DEN-ORD-JFK ($300)
SEA-DEN-ORD ($400)
SFO-DEN-ORD ($400)
LAX-DEN-ORD ($400)
DEN-ORD ($300)

Figure 11.1 The flights of a hypothetical airline

about 50 different itineraries, and each itinerary has four different booking classes, there will be about 200 itinerary-fare classes for each flight. The total number of itinerary-fare classes that are to be predicted everyday is about 73 million (365,000 × 200).

Prediction Snapshots

Each day before the departure day of a flight, a value for the demand for any itinerary-fare class on this flight can be predicted. This prediction is usually updated on a daily basis until the flight's departure day. Every prediction value obtained represents a snapshot of the predicted value of the demand. It is generally expected that the prediction will be more robust as the snapshot date is closer to the departure date of the flight. Only on the departure day of the flight, the correct value of the demand can be obtained from the air carrier's records.

Figure 11.2 depicts a graphical representation of the different snapshots recorded for a hypothetical itinerary-fare class on a flight that is departing after 180 days. As shown in the figure, the demand prediction for this itinerary-fare class starts once the flight is put for sale on the market, which is typically 9–12 months before its departure. Then, each day, a demand value is predicted to obtain a new snapshot for the demand. The values of the demand predictions obtained at each snapshot are not necessarily the same. Several factors affect the demand, and any new information regarding any of these factors might affect the predicted value of the demand. These factors are related to economic conditions, special events at the origin and destination

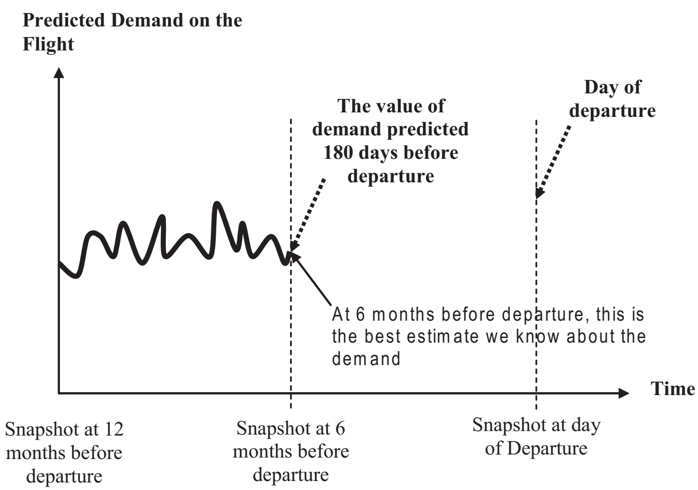


Figure 11.2 A graphical representation of the different snapshots recorded for a hypothetical itinerary-fare class on a flight that is departing after 180 days

of the flight, and so on. Figure 11.3 shows another example in which the obtained snapshots are valid until 90 days before the departure day of the flight.

It is very important that at each snapshot, the predicted value of the demand is as close to the correct value as possible because the predicted demand value at any snapshot is expected to affect the seat inventory control decisions at this snapshot. Figures 11.4 and 11.5 present two examples of erroneous demand forecasting for an itinerary-fare class for a hypothetical flight. In Figure 11.4, the value of the demand predicted at each snapshot always overestimates the actual demand value.

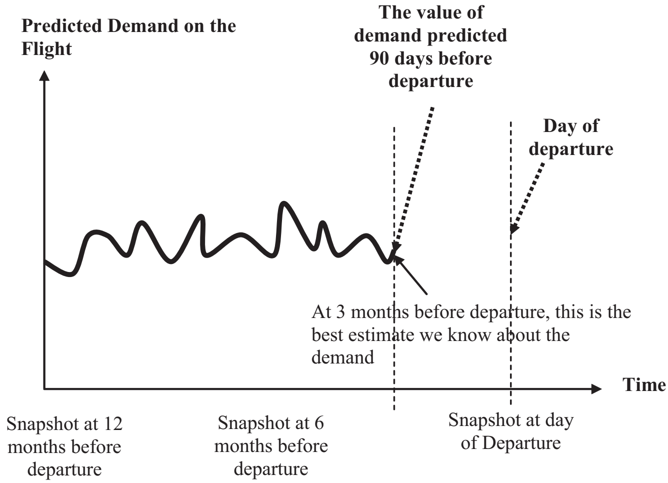


Figure 11.3 Graphical representation of the different snapshots recorded for a hypothetical itinerary-fare class on a flight that is departing after 90 days

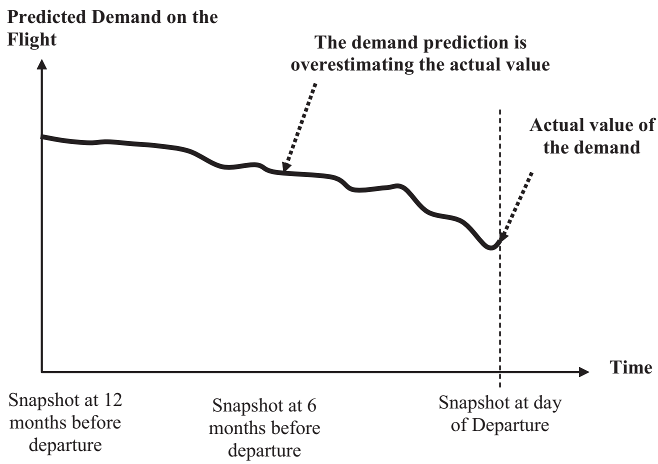


Figure 11.4 Example of overestimated snapshots

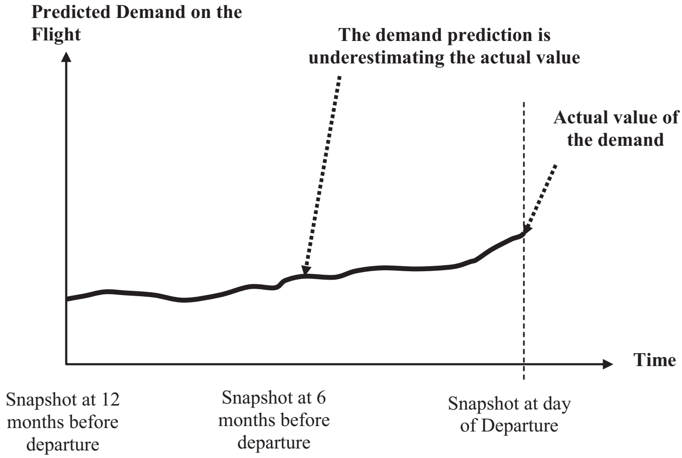


Figure 11.5 Example of underestimated snapshots

Overestimating the demand of a particular itinerary-fare class might result in overestimating the number of seats allocated for this itinerary-fare class. In Figure 11.5, the value of the demand predicted at each snapshot always underestimates the actual demand value, which might result in underestimating the number of seats allocated for this itinerary-fare class.

Factors Affecting Demand

The demand of an itinerary-fare class on any flight is affected by many factors including market size, market type (business, leisure, both), seasonality, day-of-week, time-of-day, holidays, special events, price range, air carriers competition, schedule attractiveness, and so on. For example, as the market size (that is, the population willing to travel at the origin and destination) increases, there will be a significant demand for the flights serving in this market. Also, the market type affects the expected number of travelers by season, day-of-week, and time-of-day. For example, business markets usually encounter less traffic seasonality over the year. However, weekend flights might have less traffic than weekday flights. Also, markets studies show that there is usually higher demand for late afternoon and early evening flights. However, leisure markets usually attract higher traffic during one or more season in the year. Also, flights that are close to the weekend have more traffic. Holidays and special events usually witness a boost in air carrier traffic. The fare increases and decreases in the market also impact the number of travelers, especially leisure travelers that are more elastic to airfares than business travelers. Furthermore, the level of competition and service capacity of other air carriers affects the market share of the air carrier. Finally, the attractiveness of the air carrier schedule including the departure time, arrival time, and connection convenience affect the air carrier demand.

Forecasting Methodology: Time-Series Analysis

There are several methodologies used for demand forecasting at the itinerary-fare class level including causal modeling, time-series analysis, and adaptive neural network modeling. The time-series analysis technique is the most commonly used in most RM systems. Time-series analysis is also widely used in several applications in different disciplines including transportation, water resources, weather forecasting, financial markets, census studies, and inventory estimation. Time-series analysis is used to predict a variable in the future when several historical observations about this variable are available. There are two main objectives for time-series analysis. First, it identifies the nature of the phenomenon represented by the sequence of observations. Second, it predicts future values of the time-series variable. Both of these objectives require that the pattern of observed time-series data be identified and more or less formally described.

Time-series data

As mentioned above, in time-series analysis, we are usually interested in the observations of one variable. These observations represent a sequence of measurements that follow a non-random order. Therefore, the time-series data represents consecutive measurements taken at equally spaced time intervals. Figure 11.6 illustrates an example of time-series data, which represents the quarterly average airlines OD demand in the domestic US market for the period 1998 to 2007.

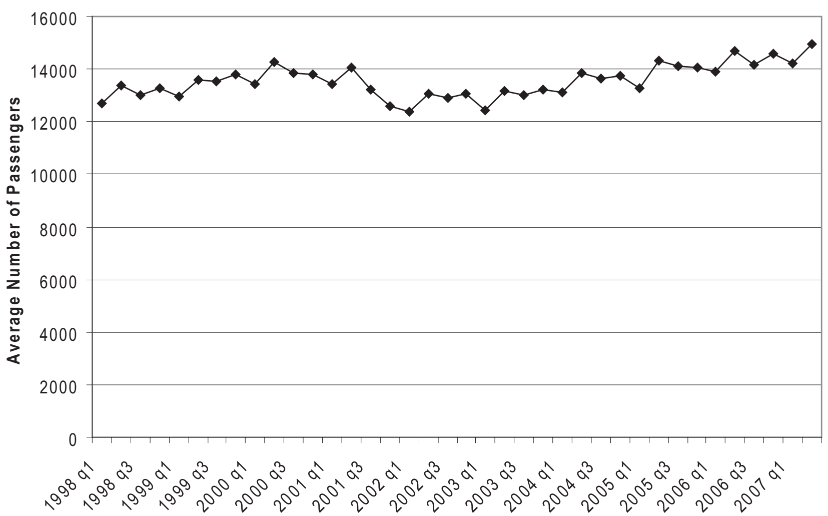


Figure 11.6 Average origin-destination demand in the domestic US market

Pattern Identification

An important step in time-series analysis is identifying the pattern of data. Two main important phenomena need to be identified in the pattern of data: the trend and seasonality. The trend represents a general systematic linear or nonlinear component that changes over time and does not repeat within the time range captured by the data. Seasonality represents a systematic pattern that repeats itself in systematic intervals over time. For example, in the time-series data presented in Figure 11.6, the plot of the successive observations indicates that there is a clear, almost linear trend that emerges, indicating that demand has a steady growth over the years (except for the period after the 9/11 terrorist attack). At the same time, the observations follow an almost identical pattern each year (for example, less people travel during the first quarter of the year), which reflects seasonality in the passenger demand.

Forecasting

Consider the variable under consideration, Y, where the most recent n historical observations are available for this variable. These observations are taken at equally spaced time intervals and denoted as Y_t, Y_{t-1}, Y_{t-2}, …, and Y_{t-n+1}, where Y_t represents the most recent observation (recorded at time t). The main objective is to estimate the next future value Y^*_{t+1} for the variable Y. Two major time-series methods are used for this purpose, which are the moving average and the exponential smoothing. In the moving average method, the next future value Y^*_{t+1} is estimated as follows:

$$Y^*_{t+1} = \frac{Y_t + Y_{t-1} + Y_{t-2} + \ldots + Y_{t-n+1}}{n} \tag{11.1}$$

The main assumption of the moving average method is that the most recent observations n are equally important in predicting the next future observation of the variable Y. Generally, there is no general rule in selecting how many observations are used in the prediction. It mainly depends on data availability and data variation over history.

The exponential smoothing method is typically used when it is believed that the most recent observations are relatively more important to forecast the future values of the variable Y than the early observations. In the exponential smoothing method, a dampening factor (1-α) is used to reduce the weight of the early observations compared to the weight of the most recent observations. The exponential smoothing method is represented mathematically as follows:

$$Y^*_{t+1} = \alpha Y_t + \alpha(1-\alpha)Y_{t-1} + \alpha(1-\alpha)^2 Y_{t-2} + \alpha(1-\alpha)^3 Y_{t-3} + \ldots \tag{11.2}$$

Where, $0 < \alpha < 1$

Application of Time-Series Analysis

Exponential smoothing is widely used within the demand forecasting component of the RM system used by major air carriers. The objective is to predict demand (number of passengers) for every itinerary-fare class in the network on a given date. To explain this further, consider the case where it is required to predict the demand of fare class B on a given itinerary between Miami (MIA) and Seattle (SEA) that is connecting at Denver (DEN), as shown in Figure 11.7. It is assumed that this itinerary is departing on Monday, October 6, 2008, and the current date is Tuesday, February 14, 2008.

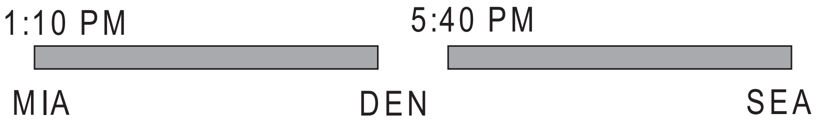


Figure 11.7 Example of a hypothetical itinerary

The historical booking data are usually stored by the air carrier for each itinerary-fare class over a period of one year. Every record in this data set represents a booking (ticket sale) by a passenger for a given itinerary-fare class that is offered by the air carrier. In this data set, it can be assumed that there is similarity in the booking trend for similar itinerary-fare classes. For instance, there is similarity in the booking trend for class B on all past itineraries MIA-DEN-SEA departing on Mondays at 1:10 PM. Accordingly, the booking observations for these similar itinerary-fare classes (that is, fare class B on the MIA-DEN-SEA itinerary departing on Monday at 1:10 PM) can be extracted from the historical booking data set and used to predict the future passenger demand on the itinerary-fare class. Over last year, there would be about 52 observations (that is, one for each week) representing bookings for fare class B on the MIA-DEN-SEA itinerary departing on Monday at 1:10 PM). These 52 observations or some of them can be used to forecast the expected number of bookings for the itinerary-fare class in the future. Figure 11.8 shows the case when the last 52 observations (February 2007–February 2008) are used to predict the number of bookings for the fare class B for the itinerary that is scheduled to depart on Monday, October 6, 2008. As time moves forward, the forecasting methodology continues to drop early observations and considers only the most recent observations, as shown in Figure 11.9.

Limitations of Time-Series Analysis

The Impact of Schedule Change

Typically over a year period, the air carrier schedule is changed and adjusted to take care of seasonality, expansion, competition, and so on. This schedule

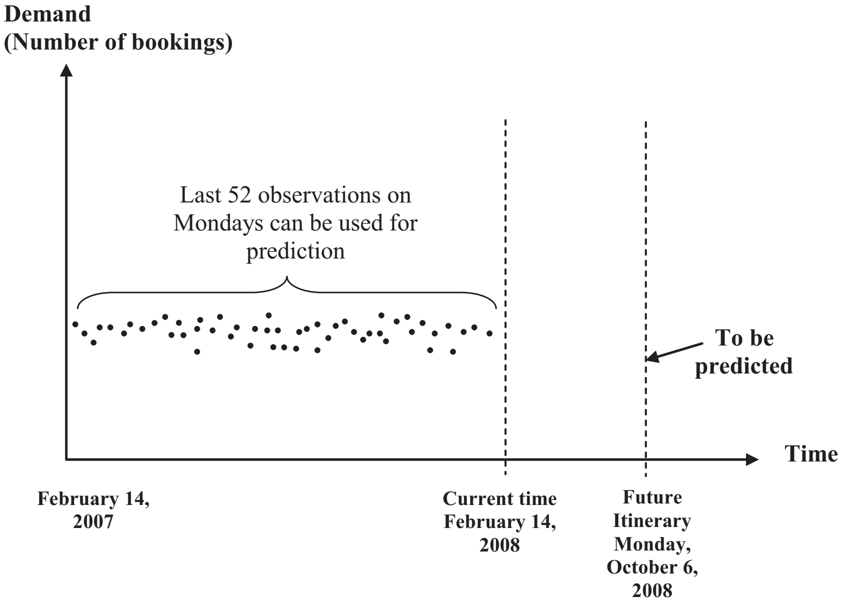


Figure 11.8 Representation of the last 52 observations of the demand

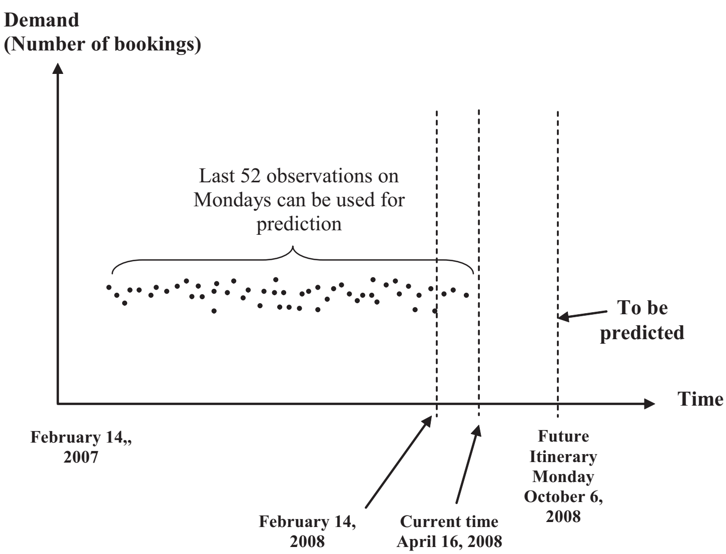


Figure 11.9 Most recent 52 observations of demand

change makes it difficult to match historical observations to predict future demand. For example, for the itinerary example given in Figure 11.7, how is demand forecasting to be performed for the itinerary MIA-DEN-SEA, when

the flight MIA-DEN is departing at 12:30 PM instead of 1:10 PM during some period of last year (Figure 11.10)? A reasonable answer for this question is that the demand for the itinerary MIA-DEN-SEA departing at 12:30 PM can be used as an approximation for the demand of MIA-DEN-SEA departing at 1:10 PM.

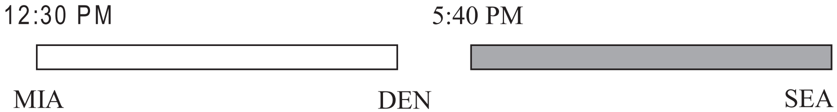
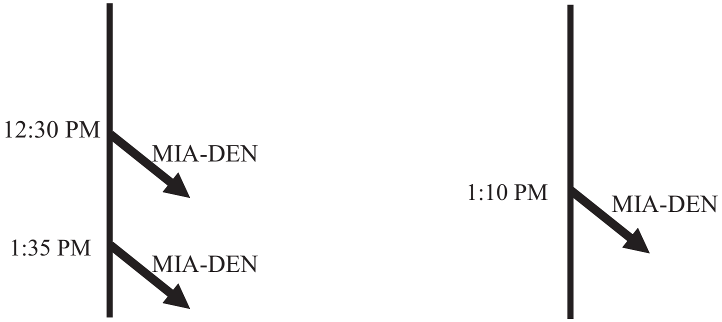


Figure 11.10 A change of the schedule of the itinerary

The Impact of Changing the Aircraft Size

In the previous example, it might be misleading to consider the demand for the itinerary MIA-DEN-SEA departing at 12:30 PM as an approximation for the demand of MIA-DEN-SEA departing at 1:10 PM. To illustrate this case, consider the example shown in Figure 11.11 that depicts the flight schedule before and after the schedule change. Before the schedule change, the air carrier used to schedule two small afternoon flights from MIA to SEA at 12:30 PM and 1:35 PM, respectively. After the schedule change, there is only one afternoon flight from MIA to DEN at 1:10 PM, which is operated by a larger aircraft. According to this scenario, one may expect that after the schedule change, the demand of the two afternoon flights from MIA to DEN that were scheduled at 12:30 PM and 1:35 PM will be accumulated on the new larger flight MIA-DEN scheduled at 1:10 PM. Therefore, the number of bookings expected on the MIA-DEN-SEA itinerary



Old Schedule: Two flights with two small aircrafts

New Schedule: One flight with larger aircraft

Figure 11.11 Example of change of aircraft size

departing at 1:10 PM will be much larger than the number of bookings on the MIA-DEN-SEA itinerary departing at 12:30 PM.

The impact of changing fare structure

Similar to schedule change and a flight's aircraft size, in many cases, the air carrier change fare structure on some of its flights responds to market competition and changes in demand over the year. This fare change is performed by changing the price levels, adding and removing booking classes, changing fare restriction rules, and so on. All these changes are expected to affect the number of bookings on each itinerary-fare class over the year period. In addition, the demand level might change because of competitors' actions such as changing the schedule, the aircraft size, and pricing, which is harder to identify.

Constrained (censored) and unconstrained demand

The demand prediction for each itinerary-fare class is used in the seat inventory control models to determine the maximum number of seats that is to be reserved for each itinerary-fare class on each flight. The historical number of bookings observed for each itinerary-fare class is controlled (constrained) by the seat protection levels for the itinerary-fare classes. In many cases, there could be additional demand that is interested in a particular itinerary-fare class on a flight, and this demand is rejected due to seat restrictions for this itinerary-fare class. This rejected demand is usually not recorded in the air carrier reservation system and not recorded among travelers that are interested in the itinerary-fare class. Ignoring this demand might underestimate the number of travelers when predicting the future demand for any itinerary-fare class.

Generally, the time series analysis has several limitations that can be summarized as follows:

- A reasonable history of the flight must be available.
- The analysis:
  - is unable to generalize forecasting information for new markets or flights for which there is no history;
  - suffers from lagged responses to demand changes and inability to adjust to rapid changes;
  - ignores useful information such as current levels of bookings, which can be used to adjust the demand prediction particularly during special events.

It is a big challenge for the air carrier to find the correct historical observations in order to predict the future demand for the different itinerary-fare classes. When this problem becomes serious for the air carrier, the demand forecasting component of the RM system implements ad hoc matching algorithms (also known as Plexing algorithms) to determine the most suitable historical observations that can be used for the prediction.

Demand as a Random Variable

A random variable is a variable whose values are random but whose statistical distribution is known. The statistical distribution gives the chances of occurrence of each possible outcome of the variable. Consider the itinerary fare class that is given in Figure 11.7 (p. 162). Assume the hypothetical situation that 15 passengers are predicted to be interested in this itinerary-fare class. The question remains as to how many passengers really book tickets for this itinerary-fare class. We cannot be sure that exactly 15 passengers would book tickets because demand is a random variable. Based on historical observations, a statistical distribution can be developed for the demand as shown in Figure 11.12. The horizontal axis in the figure gives the possible demand outcome, which can be any integer number. The vertical axis gives the chances or probability of occurrence of the corresponding demand value on the horizontal axis. According to the given statistical distribution of the demand, there is a high chance that 14, 15, or 16 passengers might book tickets for this itinerary-fare class. However, there is a low chance that zero, one, or two passengers would book tickets for this itinerary-fare class. Also, there is a very low chance that 30 passengers would book tickets for this itinerary-fare class. Usually, demand is assumed to follow a normal (bell-shaped) distribution with a defined mean and standard deviation. For the normal distribution, the mean represents the average value of the demand and also represents the most frequently occurring demand value. The standard deviation measures the dispersion or variation in the probability distribution. Figure 11.13 gives an example of two distributions with two different levels of dispersion.

Given the probability distribution of the demand, the probability of filling the seats of the flight can be estimated. For example, assume the demand probabilities that are given in Table 11.1 that correspond to the given probability distribution

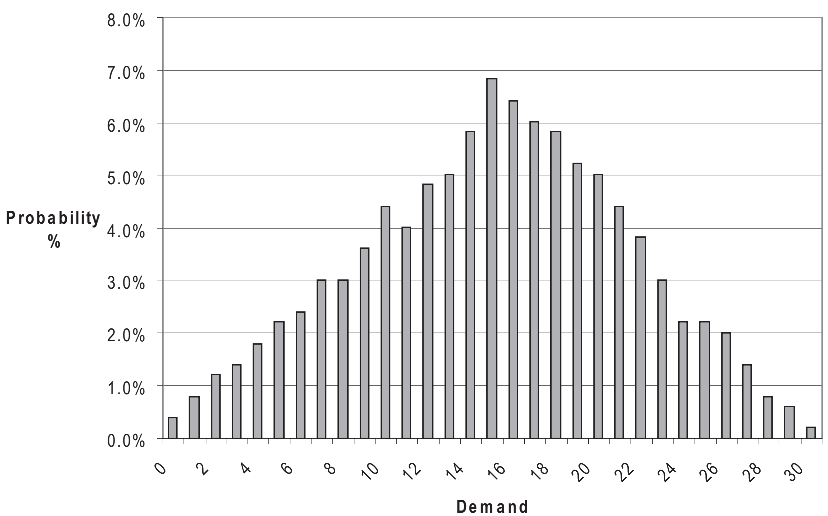
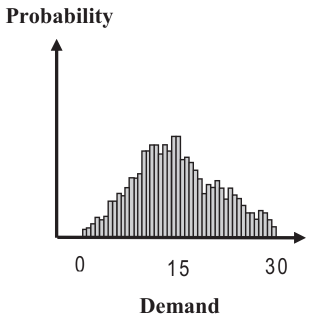
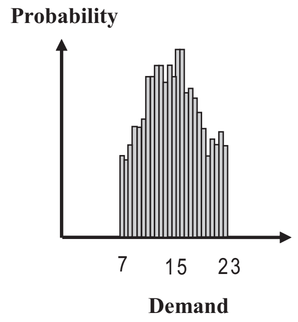


Figure 11.12 Probability distribution of demand



Distribution with high dispersion



Distribution with low dispersion

Figure 11.13 Example of two different probability distributions of demand with different levels of dispersion

Table 11.1 Probability distribution of the demand

Demand	Probability %	Demand	Probability %
0	0.402%	16	6.426%
1	0.803%	17	6.024%
2	1.205%	18	5.823%
3	1.406%	19	5.221%
4	1.807%	20	5.020%
5	2.209%	21	4.418%
6	2.410%	22	3.815%
7	3.012%	23	3.012%
8	3.012%	24	2.209%
9	3.614%	25	2.209%
10	4.418%	26	2.008%
11	4.016%	27	1.406%
12	4.819%	28	0.803%
13	5.020%	29	0.602%
14	5.823%	30	0.201%
15	6.827%		

of demand in Figure 11.12 to determine how many seats would be filled from this demand. It is not known for sure how many seats would be filled because it is not known for sure how many passengers would be interested in booking tickets for this itinerary-fare class. Because the demand has an average of about 15 passengers, there is high chance of filling about 15 seats on the flight. As the number of seats increases, the probability of filling these seats decreases. Let us calculate the probability of filing each seat of the flight:

Seat #1:

There is a high chance of filling the first seat. However, this chance is not 100 percent because the probability distribution indicates that there is a small probability (0.402 percent) that no passengers will book tickets for this itinerary-fare class. The first seat will be filled if there is at least one passenger demand. The probability of filling the first seat can be calculated as 100 percent minus 0.402 percent, which is equal to 99.598 percent. This probability means that if one seat is reserved to this demand, we will be about 99.6 percent sure that a booking will occur from this demand.

Seat #2:

The second seat will be filled if there are at least two passengers. The probability of filling the second seat can be calculated as 100 percent minus the probability of having zero passengers and having only one passenger (100 percent - 0.402 percent - 0.803 percent = 98.795 percent). This probability means that if two seats are reserved to this demand, we will be about 98.8 percent sure that two bookings will occur from this demand.

Seat #3:

Similarly, the probability of filling the third seat can be calculated as follows:

100 percent - 0.402 percent - 0.803 percent - 1.205 percent = 97.590 percent

Table 11.2 gives the probability of filling each seat out of this demand as calculated above. As given in the table, the probability of filling each additional seat decreases as the seat index increases. The probability of filling more than 30 seats out of this demand is zero because the probability distribution of the demand indicates that the probability of getting more than 30 passengers is zero.

Expected Seat Revenue

The expected revenue from each seat can be calculated by multiplying the seat fare by the probability of filling the seat by a passenger. In the previous example, we calculated the probability of filling each seat out of the demand whose probability distribution is given in Figure 11.12. Assuming that the fare of each seat is equal to $400, the expected revenue from each seat is given in Table 11.3. Figure 11.14 shows a graphical representation of the expected seat revenue. As shown in the figure, the expected revenue declines as the seat index increases. This is because

Table 11.2 Probability of filling each seat

Seat index	Probability of filling seat %	Seat index	Probability of filling seat %
1	99.6%	17	42.8%
2	98.8%	18	36.7%
3	97.6%	19	30.9%
4	96.2%	20	25.7%
5	94.4%	21	20.7%
6	92.2%	22	16.3%
7	89.8%	23	12.4%
8	86.7%	24	9.4%
9	83.7%	25	7.2%
10	80.1%	26	5.0%
11	75.7%	27	3.0%
12	71.7%	28	1.6%
13	66.9%	29	0.8%
14	61.8%	30	0.2%
15	56.0%	31	0.0%
16	49.2%	32	0.0%

Table 11.3 Expected seat revenue

Seat index	Expected seat revenue ($)	Seat index	Expected seat revenue ($)
1	$398.39	17	$171.08
2	$395.18	18	$146.99
3	$390.36	19	$123.69
4	$384.74	20	$102.81
5	$377.51	21	$82.73
6	$368.67	22	$65.06
7	$359.04	23	$49.80
8	$346.99	24	$37.75
9	$334.94	25	$28.92
10	$320.48	26	$20.08
11	$302.81	27	$12.05
12	$286.75	28	$6.43
13	$267.47	29	$3.21
14	$247.39	30	$0.80
15	$224.10	31	$0.00
16	$196.79	32	$0.00

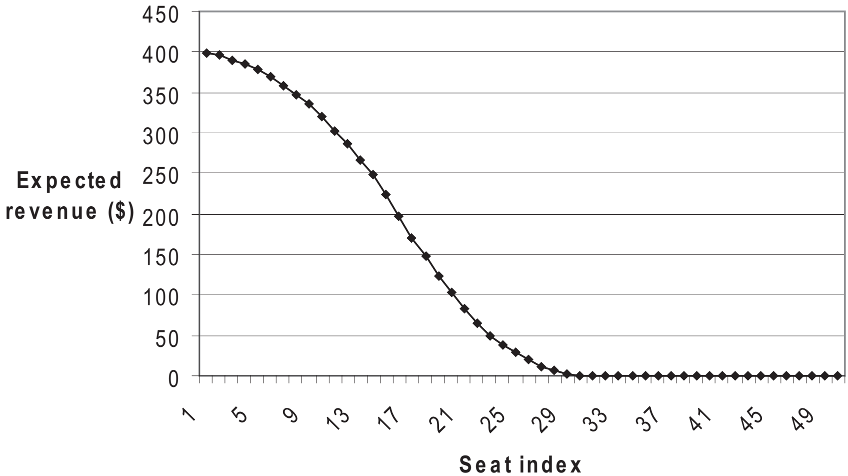


Figure 11.14 Graphical representation of the expected seat revenue

the chances of filling more seats decrease given the current level of demand for this itinerary-fare class. The expected seat revenue depends on the fare value and the chances (probability) of filling the seat. The probability of filling the seat increases only when there are more passengers that are interested in the flight. Given that the fare is unchanged, the expected seat revenue and the total revenue of the flight will increase as more demand becomes interested in the seats of the flight.

Primary Contributions

An early work on airline demand forecasting is the one by Beckman and Bobkoski (1958), who performs an analysis of some frequency distributions to fit the demand. Lee (1990) in his dissertation categorizes RM forecasting methods into three types. The study categorizes RM forecasting methods into historical booking models, advanced booking models, and combined models. The historical booking models use historical data to derive forecasts. Examples of this method include exponential smoothing methods, moving average methods, linear regression, and ARIMA time series methods. Advanced booking models use the existing booking data to forecast the demand. Examples of this method include the classical pickup, the advanced pickup, the synthetic booking curve model, and a time series of advanced booking models. Combined forecasting models include the weighted average of historical and advanced booking forecasts, regression methods, and a full information model.

As explained above, historical booking records are censored by the presence of booking and capacity limits on past flights. To deal with the censorship problem, Swan (1990) addresses the downward bias of censoring on late booking data and suggests simple statistical remedial measures. McGill (1995) develops

a multivariate multiple regression methodology for removing the effects of censorship in multiple booking classes, and describes a bootstrapping approach to test for correlations between fare class demands.

Sun et al. (1998) presents a neural network model for airline forecasting in RM. They also provide comparisons of this technique to other traditional techniques on actual airline data to show that neural network models provide significant improvements in forecast error. Belobaba and Farkas (1999) present a methodology for obtaining true demand in the airline industry. Demand is estimated by inflating actual bookings by estimates of customers turned away due to inadequate capacity and estimates of customers who decide not to book because their desired rate class is not currently for sale. Zeni (2001) in his dissertation provides an algorithm to improve forecast accuracy in airline revenue management by unconstraining demand estimates from censored data.

Fan and Chi (2004) present a cluster model for flight demand forecasting. The booking procedure is defined as a vector that grows in length when nearing departure. The results of the model show its efficiency in computational speed, robustness, and accuracy, compared with the regression model and pickup model that are popular in practice. They indicate that this model is applied successfully to Xiamen Airline of China, and the results also show a good effect in practice.

References

Beckman, M.J. and Bobkoski, F. 1958. Airline Demand: An Analysis of some Frequency Distributions. *Naval Research Logistics Quarterly*, 5, 43-51.

Belobaba, P. and Farkas, A. 1999. Yield Management Impacts on Airline Spill Estimation. *Transportation Science*, 33, (2), 217-232.

Fan, W. and Chi, H. 2004. Cluster Model for Flight Demand Forecasting. *Fifth World Congress on Intelligent Control and Automation*, 4, 3170-3173.

Lee, A.O. 1990. *Airline Reservations Forecasting: Probabilistic and Statistical Models of the Booking Process*. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA.

McGill, J. I. 1995. Censored Regression Analysis of Multiclass Demand Data Subject to Joint Capacity Constraints. *Annals of Operations Research*, 60, 209-240.

Sun, X., Brauner, E., and Hormby, S. 1998. A Large-scale Neural Network for Airline Forecasting in Revenue Management. In *Operations Research in the Airline Industry*, by Gang Yu, Springer 1998, ISBN 0792380398.

Swan, W.M. 1990. *Revenue Management Forecasting Biases*. Working Paper, Boeing Commercial Aircraft, Seattle, WA.

Zeni, R. 2001. *Improved Forecast Accuracy in Airline Revenue Management by Unconstraining Demand Estimates from Censored Data*. Ph.D. Thesis, Rutgers University, Newark, NJ, USA.

This page has been left blank intentionally

Chapter 12
No-show Rate and Overbooking

Introduction

For any flight, there is a chance that an unknown number of passengers who have confirmed bookings on the flight do not show up on the day of departure. The reason behind not showing up could be a last-minute cancellation of the travel plan, a late arrival at the airport, flight misconnections (for connecting itineraries), a last-minute itinerary change, a last-minute seat upgrade to an upper cabin, and so on. The percentage of passengers that do not show up for the flight without early notice to the airline is defined as the no-show rate. In practice, the no-show rate ranges from four to ten percent on each flight. The problem of the no-show is that the airline does not have any control on the seat that is left empty, and the airline has no chance to re-sell this seat to another customer.

The Flight Authorization Level

Due to the no-show rate, the airlines are granted to sell more seats than the physical capacity of the flight. For example, if the flight has a physical capacity of 100 seats, and it is expected that four percent of the passengers do not show up for the flight, the airline can sell about 104 tickets on the flight. As a result, about 100 passengers are expected to show up for the flight, and the flight will depart with no empty seats. This action by the airline is known as 'overbooking,' which allows the airline to sell more seats than the physical capacity of the flight to account for the no-show rate. Overbooking is not illegal, and most airlines overbook their scheduled flights to a certain level in order to compensate for 'no-shows.' When the airline overestimates the no-show rate and overbooks their flights, over sale occurs. A number of passengers that is more than the physical capacity of the flight might show up. Accordingly, some passengers have to be left behind or 'bumped' out of the flight. When an over sale occurs, passengers who are not in a hurry are offered compensation in exchange for giving up their seats voluntarily. Those passengers bumped against their will are, with a few exceptions, entitled to compensation.

Overbooking allows the airline to generate more revenue for its busy flights that rarely depart with empty seats. The actual number of seats that the airline can

sell on a flight is known as the actual capacity of the flight or the authorization level. Simply, the authorization level of the flight can be calculated as follows:

$$AU_i = \frac{C_i}{1 - NSR_i} \tag{12.1}$$

Where,

i	Flight index.
AU_i	The authorization level of the flight.
C_i	The physical capacity of the flight.
NSR_i	The estimate of the no-show rate.

In case the flight has more than one cabin class (for example, coach, business, first), the authorization level is calculated for each cabin. However, it is common practice that overbooking is performed for the coach cabin only. In this case, equation 12.1 is modified to account for passenger upgrade to higher cabin classes.

The Impact of No-show Rate Accuracy

As given in equation 12.1, the authorization level of a flight depends mainly on the estimate of the no-show rate. It is very crucial for the airline to accurately estimate the no-show rate to determine the correct authorization level for each flight. Overestimating the no-show rate would result in a serious problem of having more passengers show up for the flight than the physical capacity of the flight, and some of these passengers have to be denied boarding. However, if the no-show rate is underestimated, the authorization level is underestimated, and the flight departs with empty seats. Table 12.1 gives two examples that show the consequences of overestimating or underestimating the no-show rate for a hypothetical flight that has a physical capacity of 100 seats. In the first case, as given in Table 12.1, if the no-show rate is estimated to be 7 percent, the corresponding authorization level would be about 107 seats, which corresponds to the number of tickets to be sold to passengers. If the actual no-show rate is only 4 percent and not 7 percent, about 103 passengers would show up for the flight. Because the flight has only 100 seats, 4 passengers would be denied boarding. These passengers would be compensated by the airline and assigned to later flights. In the second case, the no-show rate is estimated to be only 1 percent, while in reality it is 4 percent. The corresponding authorization level is about 101 tickets and only 97 passengers show up for the flight. This underestimation results in having three empty seats on the flight. These empty seats represent a revenue loss for the airline, because these seats could have been filled with revenue passengers. It is observed that the no-show rate differs between flights depending on the flight type (domestic or international), the origin station (hub or spoke), the departure time, the weather conditions, and so on. Most airlines apply quantitative forecasting techniques to obtain the best estimate for the no-show rate for each future flight.

Table 12.1 The consequences of overestimating or underestimating the no-show rate

Case	The physical capacity of the flight	The actual no-show rate (practically unknown)	Estimated no-show rate	Authorization level (tickets sold)	Passengers that will show up	Overbooked seats	Empty seats
Overestimating the no-show rate	100 seats	4%	7%	100/(1-0.07)≈107	107×(1-0.04)≈103	107-103 = 4	N/A
Underestimating the no-show rate	100 seats	4%	1%	100/(1-0.0)≈101	101×(1-0.04)≈97	N/A	100-97=3

Forecasting the No-show Rate

Time-series Method

Time-series forecasting is the use of a model to forecast future events based on known historical events. A time series is a sequence of data points, measured typically at successive times, spaced uniform time intervals. For each future flight in the schedule, the history of the no-show rate is recorded for similar flights in the past. Historical similar flights are identified to have same origin, destination, departure time, day-of-week, and so on. Historical no-show data are then used to forecast the no-show rate for the future flight under consideration.

Cluster Analysis Method

Cluster analysis is another technique used to estimate the expected no-show rate for future flights in the schedule. Cluster analysis is an exploratory data analysis tool which aims at sorting different objects into groups in a way that the degree of association between two objects is maximal if they belong to the same group and minimal otherwise. A considerable number of historical flights is grouped into different clusters based on their main characteristics such as origin, destination, departure time, arrival time, day-of-week, seasonality, and so on. Then, the average no-show rate for each cluster is estimated. Next, for any future flight, and based on its characteristics, the cluster that could include this flight is identified. The no-show rate for this flight is assumed to equal the average no-show rate of the identified cluster.

Economic Method

There have been other methods to calculate the authorization level that are based on balancing the risk of having overbooked or under-booked flights, which is known as the economic method. This method balances the cost of having a flight

with an empty seat against the cost of denying boarding for a passenger. It requires the following information:

- the revenue loss due to an empty seat;
- the cost of reaccommodating a denied passenger;
- The ill-will cost (bumped passenger might not select to fly on the air carrier in the future).

The problem is formulated similar to a famous operations research problem known as the newsvendor problem, which tries to estimate the number of newspapers to be stocked for sale. It seeks to equalize the risk of having overstock and under-stock. It is also known as a single period inventory problem with a variable demand. The method can be described as follows:

Consider,

S	The physical capacity of the flight.
Y	The number of tickets to be overbooked (the unknown).
$S + Y$	The total number of tickets to be sold.
x	The number of no-shows.
$f(x)$	The probability distribution of the no-shows (for example, Figure 12.1).
C	The average overbooking cost (reaccommodation cost plus ill-will cost).
B	The underage cost (the opportunity cost of flying an empty seats, which is the ticket price).

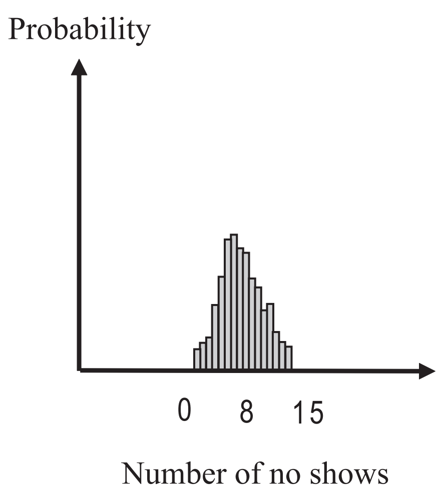


Figure 12.1 An example of the probability distribution of the number of no-shows for a hypothetical flight

Two scenarios could happen. The first scenario represents the case where the number of overbooked tickets is greater than the number of no-shows (that is, $x < Y$). In this case, there is a number of passengers to be bumped from the flight that is equal to $(Y - x)$, and each bumped passenger costs \$$C$. The second scenario represents the case where the number of overbooked tickets is less than the number of no-shows (that is, $x > Y$). In this case, there is a number of empty seats on the flight that is equal to $(x - Y)$, and each empty seat could have generated a revenue of \$$B$.

To determine the optimal number of tickets to be overbooked, the cost of having overbooked passengers is balanced with the revenue loss due to having empty seats, which can be expressed mathematically as follows:

$$\text{C} \times \text{Probability(Overbooking)} = \text{B} \times \text{Probability(Having empty seats)} \tag{12.2}$$

The probability of overbooking is equivalent to the probability that $x < Y$ and the probability of having empty seats is equivalent to the probability that $x > Y$. Accordingly, equation (12.2) can be rewritten as follows:

$$C \times \Pr(x < Y) = B \times \Pr(x > Y) \tag{12.3}$$

According to the probability laws,

$$\Pr(x > Y = 1 - \Pr(x < Y) \tag{12.4}$$

Substituting from (12.4) into (12.3),

$$C \times \Pr(x < Y) = B \times (1 - \Pr(x < Y) = B - B \times \Pr(x < Y) \tag{12.5}$$

$$\Pr(x < Y) = \frac{B}{B + C} \tag{12.6}$$

According to equation (12.6), given the values of B and C as well as the main characteristics of the probability distribution of x, the value of Y can be determined. Figure 12.2 shows the relationship between the probability distribution of the number of no-shows and the number of overbooked tickets. In this figure, the number of overbooked tickets (Y) is determined such that the dotted area is equal to ${}^{B}\!/_{B+C}$.

It should be noted that while this method has gained acceptance in the literature, it has not been applied by airlines because it is difficult to the estimate the correct vales of B and C, which are expected to be different among flights. Therefore, most airlines apply a modified version of the synthesized method given in equation 12.1 above.

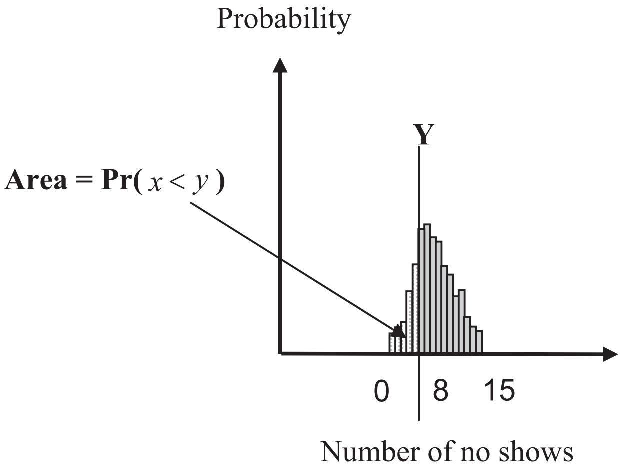


Figure 12.2 The relationship between the probability distribution of the number of no-shows and the number of overbooked tickets

Primary Contributions

The literature of the airline overbooking is very rich. This problem was thoroughly investigated as early as the 1960s. An example of this early work is by Beckmann (1958), Thompson (1961), and Taylor (1962). Beckman (1958) models booking and no-shows in an effort to find an optimal overbooking strategy. Thompson (1961) describes a methodology to overcome the no-show problem by allowing controlled overbooking to be made and relying on subsequent cancellations to keep the number of bookings left at departure time at, or just below, the seating capacity of the flight. Taylor (1962) develops a method to allow for cancellations and group reservations.

Rothstein (1971) was the first to formulate airline overbooking as a dynamic programming problem, however, this approach is computationally intractable due to the curse of dimensionality. Extensions to this work are given in Rothstein (1974, 1985). Bodily and Pfeifer (1973) consider the probability of a customer cancellation. They conclude that a customer cancellation depends on when the reservation is made and the unknown events that might occur before departure. The drawback of this approach is that it does not consider the dynamic nature inherent in the reservation process. Vickrey (1972) suggests that the oversold conditions could be resolved with auctions. He also presents a conceptual description of a multiple fare class reservations system similar to those now in widespread use. Ladany (1976) derives models for the overbooking decision process in combination with the pricing decision. Williams (1977) demonstrates the necessity of applying optimization methods to the overbooking and pricing problems as opposed to using simple, approximate decision rules. Nagarajan (1979) presents an auction solution to the problem of airline overbooking. Belobaba (1987) addresses the

problem of overbooking in multiple fare classes. A heuristic approach to solving the problem is presented. Alstrup et al. (1989) develop a dynamic programming approach to solve the overbooking problem for two fare classes. They assume that customers request and cancel reservations in groups of five, thereby reducing the size of the problem. Chatwin (1993), in his dissertation, considers overbooking models with discrete time and discrete state spaces, and a continuous time birth and death process. Subramanian et al. (1999) formulate the overbooking revenue management problem as a finite-horizon, discrete-time Markov decision process (MDP). Chatwin (1998) solves the multi-period overbooking problem that relates to a single flight leg and service class. Karaesmen and van Ryzin (2004) formulate the overbooking problem as a two-period optimization problem. In the first period, given the probabilistic knowledge of cancellations, reservations are accepted. In the second period, cancellations are realized and surviving customers are assigned to the various inventory classes to minimize penalties.

Suzuki (2002) presents an empirical analysis of the optimal overbooking policies for major US airlines. Garrow and Koppelman (2004) model airline passengers' no-show and standby behavior using passenger and outbound and inbound itinerary information. They describe factors that influence no-show and standby behavior in the continental US markets using multinomial and nested logic models. Recently, Klophaus and Pölt (2007) analyze the revenue potential of incorporating the passengers' Willingness to Pay (WTP) into airline overbooking. They examine the time-dependent spoilage costs as an extension of the static overbooking model. The simulation results with real airline data indicate that considering the dynamics of passengers' WTP in the overbooking decision leads to consistent gains in net revenue.

References

Alstrup, J., Andersson, S.E., Boas, S., and Madsen, O. 1989. Booking Control Increases Profit at Scandinavian Airlines. *Interfaces*, 19, 10-19.

Beckman, J.M. 1958. Decision and Team Problems in Airline Reservations. *Econometrica*, 26, 134-145.

Belobaba, P. 1987. Airline Yield Management: An Overview of Seat Inventory Control, *Transportation Science*, 21, 63-73.

Bodily, S.E. and Pfeifer, P.E. 1973. Overbooking Decision Rules. *OMEGA 20*, 129-133.

Chatwin, R.E. 1993. *Optimal Airline Overbooking*. PhD thesis, Stanford University.

Chatwin, R.E. 1998 Multiperiod Airline Overbooking with a Single Fare Class. *Operations Research*, 46, 805-819.

Garrow, L.A. and Koppelman, F.S. 2004. Multinomial and Nested Logit Models of Airline Passengers' No-Show and Standby Behavior. *Journal of Revenue and Pricing Management*, 3(3), 237-253.

Ladany, S. 1976. Dynamic Operating Rules for Motel Reservations. *Decision Sciences*, 7, 829-840.

Karaesmen, I. and van Ryzin, G. 2004. Overbooking with Substitutable Inventory Classes. *Operations Research*, 52, 83-104.

Klophaus, R. and Pölt, S. 2007. Airline Overbooking with Dynamic Spoilage Costs. *Journal of Revenue and Pricing Management*, 6(1), 9-19.

Nagarajan, K.V., 1979. On an Auction Solution to the Problem of Airline Overbooking. *Transportation Research* 13A, 111-114.

Rothstein, M. 1971. An Airline Overbooking Model. *Transportation Science,* 5, 180-192.

Rothstein, M. 1974. Hotel Overbooking as a Markovian Sequential Decision Process. *Decision Sciences*, 5, 389-394.

Rothstein, M. 1985. OR and the Airline Overbooking Problem. *Operations Research*, 33(2), 237-248.

Subramanian, J., Stidham, S., and Lautenbacher, C.J. 1999. Airline Yield Management with Overbooking, Cancellations, and No-Shows. *Transportation Science*, 33, 147-167.

Suzuki, Y. 2002. An Empirical Analysis of the Optimal Overbooking Policies for U.S. Major Airlines. *Transportation Research E*, 38, 135-149.

Taylor, C.J. 1962. The Determination of Passenger Booking Levels. *AGIFORS Symposium Proceedings*, 2, Fregene, Italy.

Thompson, H. 1961 Statistical Problems in Airline Reservation Control. *Operational Research Quarterly*, 12, 167-185.

Vickrey, W. 1972. Airline Overbooking: Some Further Solutions. *Journal of Transport Economics and Policy*, 6, 257-270.

Williams, F. 1977. Decision Theory and the Innkeeper: An Approach for Setting Hotel Reservation Policy. *Interfaces*, 7, 18-30.

Chapter 13
Seat Inventory Control for Flight-based Revenue Management Systems

Introduction

Flight-based RM systems are typically used by air carriers that adopt a point-to-point network structure. As mentioned earlier, the point-to-point network structure is when an air carrier focuses on severing local traffic between city-pairs and less attention is given to the connecting traffic to the beyond destinations. Thus, most passengers on the flight are traveling from the origin to the destination of the flight. The seats on each flight are classified to different booking (fare) classes that reflect the needs and characteristics of travelers that are related to their willingness to pay, a Saturday-stay, an advance purchase, the refund policy, and so on. For any flight in the point-to-point network structure, the objective of the seat inventory control is to determine the number of seats to be allocated for each booking class considered for the flight. This chapter presents the mechanism of the seat inventory control for flight-based RM systems. A simplified case of two fare classes is explained. Then, a more comprehensive case with multiple fare classes is presented.

Two Booking Classes

Consider a simplified case where there is a flight with a seat capacity $C = 120$ within a point-to-point network structure. Assume also that there are two fare classes considered for this flight. The first class has a fare, $f_1 = \$400$, which is a full fare that is typically charged for business travelers and has limited restrictions. The second class has a fare, $f_2 = \$225$, which is a discounted fare that is allocated to leisure travelers. Booking this discounted fare necessitates a Saturday-stay at the destination and three-weeks advance purchase. Assume the statistical distribution for the demand of the first fare class as shown in Figure 13.1. The demand as shown in this figure ranges from 0 to 30 passengers with an average of about 15 passengers. It is assumed that the demand for the second fare class is strong and widely available. It is also assumed that any seat that is allocated to this fare class is a guaranteed sale. The objective is to determine the number of seats for each fare class.

Because the demand for the discounted fare class is widely available, any seat that is allocated to this booking class would be sold and generate a revenue of \$225. If we know, for example, that 17 passengers are interested in the high-fare

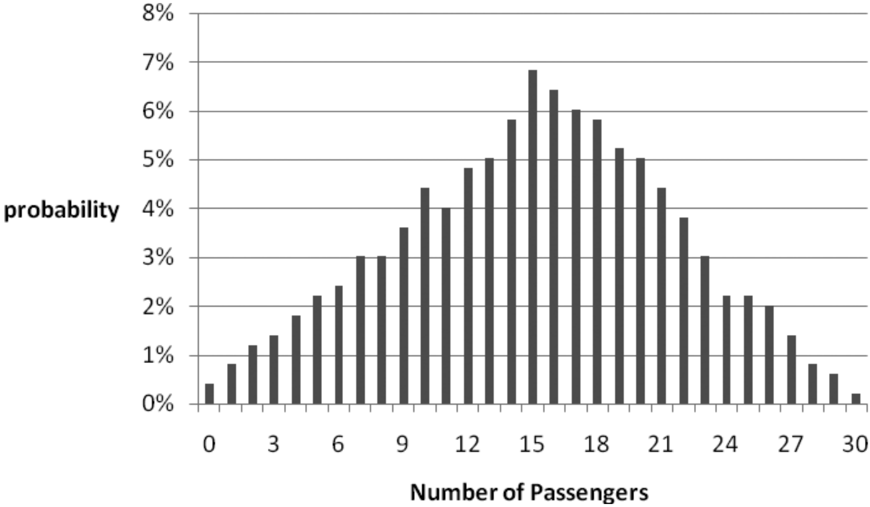


Figure 13.1 Example of the probability distribution for the demand of the high-fare class

class, we should protect 17 seats for those passengers because every sale generates a high revenue of \$400. Because the actual number of high-revenue passengers is not known, another mechanism should be implemented to determine the number of seats to be protected for the high-fare class.

Littlewood (1972) suggests that the total flight revenue would be maximized by closing down the sale of the low-fare class when the certain or guaranteed revenue from selling another low-fare seat is exceeded by the expected revenue of selling the same seat at a higher fare. Littlewood has several assumptions for his formulation. First, it is assumed that low-fare passengers always book first. Second, there are no cancellations of bookings. Also, a rejected booking request is revenue lost by the air carrier.

In Chapter 11, we explained how to calculate the expected seat revenue for a given demand stream. The expected seat revenue is calculated by multiplying the probability of filling the seat of this demand stream by the fare. Table 13.1 gives the expected seat revenue out of the high-revenue demand for the first 32 seats of the aircraft, which is also presented graphically in Figure 13.2.

According to the expected seat revenue given in Table 13.1, at least 14 seats have to be protected for the high-fare class. These seats are defined as the seats to be protected for the high-fare class from the passengers that are only interested in the tickets for the low-fare class denoted as S_2^1. Accordingly, a maximum of 106 seats (120 minus 14) can be allocated for the low-fare class. The same conclusion can be obtained through Figure 13.2, where a horizontal line is drawn at the fare amount of the low fare class (that is, \$225) until this line meets the curve of the

Table 13.1 Expected seat revenue

Seat index	Expected seat revenue ($)	Seat index	Expected seat revenue ($)
1	$398.39	17	$171.08
2	$395.18	18	$146.99
3	$390.36	19	$123.69
4	$384.74	20	$102.81
5	$377.51	21	$82.73
6	$368.67	22	$65.06
7	$359.04	23	$49.80
8	$346.99	24	$37.75
9	$334.94	25	$28.92
10	$320.48	26	$20.08
11	$302.81	27	$12.05
12	$286.75	28	$6.43
13	$267.47	29	$3.21
14	$247.39	30	$0.80
15	$224.10	31	$0.00
16	$196.79	32	$0.00

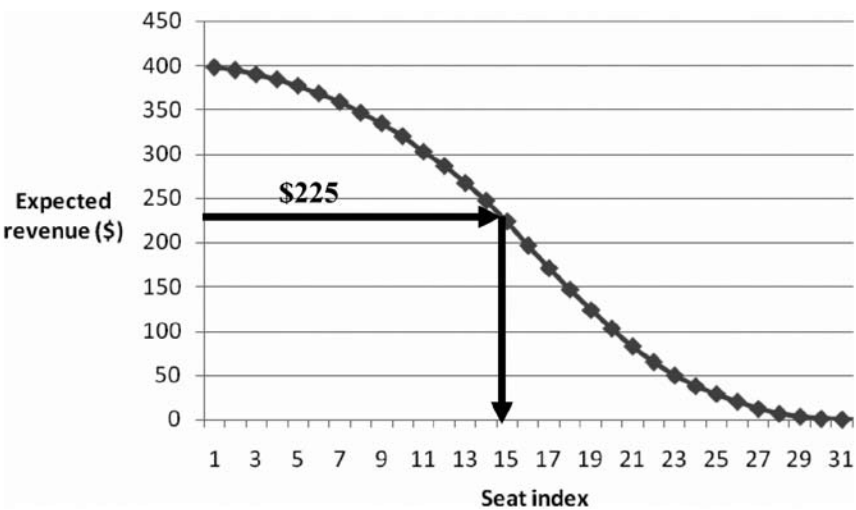


Figure 13.2 Graphical representation of the expected seat revenue

expected seat revenue. At the intersection point, a vertical line is drawn to get the corresponding seat index for the protected seats (=14 seats).

Because the demand for the high-fare class is random, it could happen that more than 14 passengers request bookings for this fare-class. If there are no seats available on the flight, logically these additional booking requests for the high-fare class are rejected. However, if there are any seats available in the lower-fare class, these seats have to be allocated and sold for this additional demand. This allocation is done through serial nesting as explained in Chapter 10. According to nesting, the booking limit, BL_2, (the authorization level) of the low-fare class is 106 seats, and the booking limit of the high-fare class, BL_1, is 120 seats (the whole seat capacity).

Multiple Booking Classes

Belobaba (1987) extends the work of Littlewood to consider the case when multiple fare classes are proposed on each flight, and it is required to determine the booking limit for each fare class. Belobaba introduces the term $EMSR_i$, which is the expected marginal seat revenue for booking class *i*. The expected marginal seat revenue for a booking class is the expected revenue generated from an additional seat to be sold for this booking class. For the example given in Table 13.1, there are 14 seats to be protected for the high-fare booking class. The expected marginal seat revenue for this booking class, $EMSR_i$, is equivalent to the expected seat revenue of the 15th seat ($224.10). In the example above, it is concluded that we should keep accepting more passengers in the low-fare class (booking class 2) as long as:

$$f_2 \geq EMSR_1\,(S_2^1) \tag{13.1}$$

Where:

S_2^1 = The number seats to be protected for the high-fare class (booking class 1).

f_2 = The fare for booking class 2.

When there are multiple booking classes on a flight, there are multiple demand streams, each of which corresponds to one of these booking classes. The objective is to determine the number of seats to be protected for each booking class.

Consider a hypothetical flight in a point-to-point network structure that has three different booking classes denoted as BC1, BC2, and BC3. The fare values of these booking classes are $300, $200, and $100, respectively. It is assumed that low-fare bookings always come first. Assume also that there are three demand streams where each demand stream is interested in only one booking class. The statistical distribution of the demand for BC1, BC2, and BC3, is

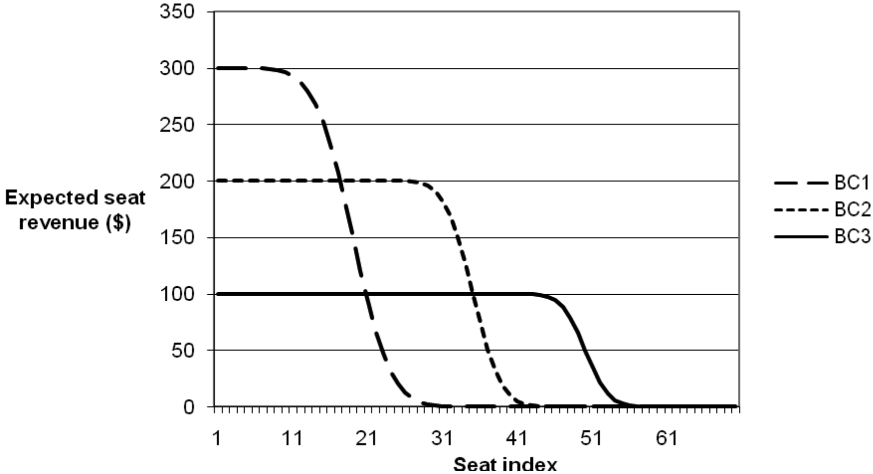


Figure 13.3 Expected seat revenue for each demand stream

given to be normally distributed, where the average number of passengers interested in BC1, BC2, and BC3 are 19, 35, and 50 passengers, respectively. Figure 13.3 gives the expected seat revenue for each demand stream. As shown in the figure, each seat generates a different revenue amount when it is allocated to the different booking classes. To maximize the flight revenue, the seat should be allocated to the booking class that generates the highest revenue. According to the figure, at least 17 seats should be protected for the first booking class (BC1) that generates the highest possible revenue for these seats. The number of protected seats for BC1 is obtained at the intersection point of the expected seat revenue curve of BC1 with its next lower expected seat revenue curve. Next, at least 35 seats need to be protected for BC2. This number is obtained at the intersection point of the expected seat revenue curve of BC2 with its next lower expected seat revenue curve. Finally, the remaining seats of the flight can be sold for the lowest booking class, BC3.

If the flight capacity is 120 seats, the booking limit of BC3 is 68 seats (120 minus 35). By applying sequential nesting of the different booking classes, the booking limit of BC2 is 103 seats (120 minus 52). Finally, the booking limit of BC1 is 120 seats (the whole seat capacity).

Generally, for multiple fare classes on a single-leg flight, more comparisons of expected seat revenue curves are needed among the relative classes offered. The total number of comparisons required for *k* nested booking classes is given by:

$$numberof\ comparisons = \frac{k(k-1)}{2} \tag{13.2}$$

For the case of two booking classes (that is, k = 2), the number of comparisons is equal to one as shown in Figure 13.2 on page 183. When k = 3, the number of

comparisons is equal to 3. In the example shown in Figure 13.3, the third comparison between BC1 and BC3 is ignored as it does not affect the seat inventory control decisions of the flight.

The seat inventory control is usually implemented in a dynamic context in which booking limits are revised on a regular basis as it gets closer to the day of departure of the flight. As the departure day nears, more information becomes available that is related to available booking trends and expected future booking requests. In the dynamic seat inventory control, estimates of future booking requests at various times before departure are required to calculate the optimal protection levels for the unbooked seats on the flight. At each period, the available capacity is updated by excluding the booked seats.

Primary Contributions

Flight-based seat inventory control problems were first solved by Littlewood's rule (Littlewood 1972), as explained above. He concludes that discount fare bookings should be accepted as long as their revenue value exceeds the expected revenue of future full-fare bookings. This problem is solved for BOAC (now British Airways), which offered early-bird bookings that charged lower fares to passengers who booked at least 21 days in advance of flight departure. Other work in the same line includes Richter (1982). Mayer (1976) studies the performance of Littlewood's rule using simulation. Titze and Griesshaber (1983) offer additional simulation evidence that Littlewood's rule is robust and claim that overlapping demand arrivals between classes would not affect the capacity allocation rule.

Belobaba (1987) extends Littlewood's rule to multiple nested fare classes and introduces the term Expected Marginal Seat Revenue (EMSR) for the general approach. This method is known as the EMSRa method, which produces nested protection levels for the different booking classes on the flight. The EMSRa method does not, however, yield optimal booking limits when more than two fare classes are considered. Belobaba (1989) describes how the EMSRa heuristic can be used on a dynamic problem even though the EMSRa decision rule is a static rule. McGill (1989) and Wollmer (1992) show that EMSR provides realistic approximations with typical airline demand distributions. A later refinement of EMSR, called EMSRb, apparently produces better approximations to optimal booking limits and has been widely implemented. This method is based on the idea of equating the marginal revenues in the various fare classes. The EMSRa and EMSRb methods became most popular for single-leg problems. They are commonly used under the assumption that the demand for each fare class is independent and normally distributed. An extension and related work for different types of distributions or dependencies may be found in Brumelle et al., (1990), Curry (1990), Wollmer (1992), Brumelle and McGill (1993), Robinson (1995), Belobaba and Weatherford (1996), Van Ryzin and McGill (1998), Van Ryzin and McGill (2000), Talluri and Van Ryzin (199), and Weatherford and Belobaba (2002). Other models including

Chatwin (1996) and Chatwin (1999) are developed to include important practical issues such as overbooking, cancellations, no-shows, and trade-ups.

Another approach to treat the single-leg RM problem is to formulate the problem as a dynamic programming model. An early version of this approach is by Lee and Hersh (1993), who develop a model for dynamic airline seat inventory control with multiple seat bookings. This model is used the most in the literature, and thus many approximation algorithms and extensions are made for this model. Kleywegt and Papastavrou (1998) demonstrate that the problem can also be formulated as a Dynamic and Stochastic Knapsack Problem (DSKP). Subramanian et al. (1999) extend the work of Lee and Hersh to incorporate cancellations, no-shows, and overbooking. Liang (1999), and Van Slyke and Young (2000) solve the Lee and Hersh model in continuous-time. Lautenbacher and Stidham (1999) formulate a model that includes the static and dynamic models as special cases. Zhao and Zheng (2001) present a dynamic programming problem with two classes that incorporate trade-ups.

References

Belobaba, P. 1987. Airline Yield Management: An Overview of Seat Inventory Control. *Transportation Science*, 21, 63-73.

Belobaba, P. 1989. Application of a Probabilistic Decision Model to Airline Seat Inventory Control. *Operations Research*, 37(2), 183-196.

Belobaba, P. and Weatherford, L. 1996. Comparing Decision Rules that Incorporate Customer Diversion in Perishable Asset Revenue Management Situations. *Decision Sciences*, 27, 343-363.

Brumelle, S.L., McGill, J.I. ,Oum, T.H., Sawaki, K., and Tretheway, M.W. 1990. Allocation of Airline Seats between Stochastically Dependent Demands. *Transportation Science*, 24, 183-192.

Brumelle, S.L. and McGill, J.I. 1993. Airline Seat Allocation with Multiple Nested Fare Classes. *Operations Research*, 41, 127-137.

Chatwin, R. 1996. Optimal Control of Continuous-Time Terminal-Value Birth-and-Death Processes and Airline Overbooking. *Naval Research Logistics*, 43, 159-168.

Chatwin, R. 1999. Continuous-Time Airline Overbooking with Time-Dependent Fares and Refunds. *Transportation Science*, 33(2), 182-191.

Curry, R.E. 1990. Optimal Airline Seat Allocation with Fare Classes Nested by Origins and Destinations, *Transportation Science*, 24, 193-204.

Kleywegt, A. and Papastavrou, J. 1998. Acceptance and Dispatching Policies for a Distribution Problem. *Transportation Science*, 32(2), 127-141.

Lautenbacher, C.J. and Stidham jr., S. 1999. The Underlying Markov Decision Process in the Single-Leg Airline Yield Management Problem. *Transportation Science*, 33, 136-146.

Lee, T.C. and Hersh, M. 1993. A Model for Dynamic Airline Seat Inventory Control with Multiple Seat Bookings. *Transportation Science*, 27, 252-265.

Liang, Y. 1999. Solution to the Continuous Time Dynamic Yield Management Model. *Transportation Science*, 33(1), 117-123.

Littlewood, K. 1972. Forecasting and Control of Passenger Bookings. In *AGIFORS Symposium Proc.* 12, 95-117.

Mayer, M. 1976. Seat Allocation, or a Simple Model of Seat Allocation via Sophisticated Ones. *AGIFORS Symposium Proc.* 16.

McGill, J. 1989. *Optimization and Estimation Problems in Airline Yield Management*, Ph.D. thesis, Faculty of Commerce and Business Administration, University of British Columbia, Vancouver, BC.

Richter, H. 1982. The Differential Revenue Method to Determine Optimal Seat Allotments by Fare Type. *AGIFORS Symposium Proc.* 22.

Robinson, L.W. 1995. Optimal and Approximate Control Policies for Airline Booking with Sequential Nonmonotonic Fare Classes. *Operations Research*, 43, 252-263.

Subramanian, J., Stidham jr., S., and Lautenbacher, C.J. 1999. Airline Yield Management with Overbooking, Cancellations, and No-Shows. *Transportation Science*, 33, 147-167.

Talluri, K. and van Ryzin, G.J. 1999. A Randomized Linear Programming Method for Computing Network Bid Prices. *Transportation Science,* 33, 207-216.

Titze, B. and Griesshaber, R. 1983. Realistic Passenger Booking Behaviours and the Simple Low-Fare/High-Fare Seat Allotment Model. *AGIFORS Symposium Proc.* 23.

Van Ryzin, G.J. and McGill, J.I. 2000, Revenue Management without Forecasting or Optimization: An Adaptive Algorithm for Determining Airline Seat Protection Levels. *Management Science*, 46, 760-775.

Van Slyke, R. and Young, Y. 2000. Finite Stochastic Knapsacks with Applications to Yield Management. *Operations Research*, 48(1), 155-172.

Weatherford, L.R. and Belobaba, P. 2002. Revenue Impacts of Fare Input and Demand Forecast Accuracy in Airline Yield Management. *Journal of the Operational Research Society*, 53, 811-821.

Wollmer, R. 1992. An Airline Seat Management Model for a Single Leg Route When Lower Fare Classes Book First. Operations Research, 40(1), 26-37.

Zhao, W. and Zheng, Y.S. 2001. A Dynamic Model for Airline Seat Allocation with Passenger Diversion and No-shows. *Transportation Science*, 35, 80-98.

Chapter 14
Seat Inventory Control for Network-based Revenue Management Systems

Introduction

The complexity of the seat inventory control problem increases significantly for air carriers that adopt the hub-and-spoke network structure. Every flight in the hub-and-spoke network structure serves both local travelers between the origin and the destination of the flight, and connecting passengers to and from other destinations. Therefore, passengers on the flight are composed of a mix of traveler groups that differ in their origin and destination, as well as their trip purpose (business or leisure) and travel needs. For example, for the air carrier network structure shown in Figure 14.1, the flight DEN-ORD could be serving as many as seven different itineraries as given in Table 14.1. Each of these itineraries could have several booking classes with different fare values to match the differences in the passengers' characteristics and preferences that are related to their willingness to pay, an advance purchase, a Saturday stay, and so on. The main objective of the seat inventory control is to guarantee the seat availability for booking requests for the high-revenue itineraries to maximize the overall revenue of the schedule.

The Network Effect

It should be clear that for the seat inventory control for the hub-and-spoke network structure, accepting any booking request on any flight affects the decision to accept or reject other booking requests on other flights. To elaborate this concept, we again use the example of the network structure given in Chapter 11, as shown in Figure 14.1. Assume the hypothetical case that there is one seat left on the flight DEN-ORD. If demand forecasting indicates that there is a passenger interested in the itinerary LAX-DEN-ORD-LHR, the seat on flight DEN-ORD might be reserved for this itinerary as it contributes the high revenue of $1,300.

Assume also that there is one seat left on flight LAX-DEN, and it is predicted that there is one passenger interested in the itinerary LAX-DEN-FRA at $1,500. In this case, the last seat on flight LAX-DEN might be reserved to this passenger, since it is contributing $1,500 to the air carrier's revenue. By doing so, the expected passenger on the itinerary LAX-DEN-ORD-LHR will not be accommodated by the air carrier, because the last seat on flight LAX-DEN is reserved for the LAX-DEN-FRA passengers. Accordingly, in this case, the last seat on flight DEN-ORD

Table 14.1 The different itineraries served by flight DEN-ORD

SEA-DEN-ORD-ATL ($400) SFO-DEN-ORD-ATL ($400) LAX-DEN-ORD-ATL ($400) DEN-ORD-ATL ($300) SEA-DEN-ORD-DTW ($400) SFO-DEN-ORD-DTW ($400) LAX-DEN-ORD-DTW ($400) DEN-ORD-DTW ($300)	SEA-DEN-ORD-LHR ($1200) SFO-DEN-ORD-LHR ($1250) LAX-DEN-ORD-LHR ($1300) DEN-ORD-LHR ($1000) SEA-DEN-ORD-JFK ($400) SFO-DEN-ORD-JFK ($400) LAX-DEN-ORD-JFK ($400) DEN-ORD-JFK ($300)	SEA-DEN-ORD ($400) SFO-DEN-ORD ($400) LAX-DEN-ORD ($400) DEN-ORD ($300)

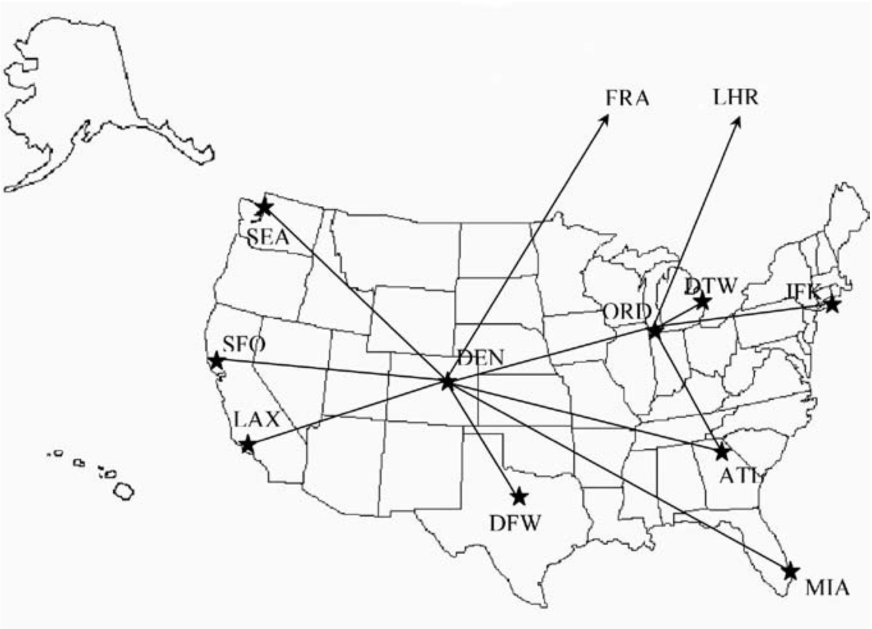


Figure 14.1 An example of a hub-and-spoke network structure

will be unoccupied. This example shows that the seat inventory control on flight LAX-DEN affects the seat inventory control on flight DEN-ORD.

In the above example, it is easy to decide to accept the booking request of the itinerary LAX-DEN-FRA at $1,500, as it generates more revenue to the air carrier. However, as the network gets larger, a robust methodology is needed to determine the number of bookings to accept from each itinerary-fare class.

Another problem to consider in the seat inventory control for the hub-and-spoke network is how to deal with the stochastic (random) nature of the demand, and how the seat inventory mechanism should respond to an extra unexpected demand for one or more itinerary-fare classes. In the flight-based seat inventory control, to consider the stochastic nature of the demand, the different booking classes of each flight are nested sequentially according to their revenue contribution to this flight

only. Accordingly, seats are always available for a high-fare booking class as long as there are seats available in any of the lower booking classes. For the network-based seat inventory control, a similar nesting mechanism should be developed to avoid the scenario of rejecting extra unexpected high-revenue passengers while there are seats available on flights. In this nesting mechanism, the itinerary-fare classes should be nested on each flight according to their contribution to the overall network revenue.

Mathematical Formulation

To determine the number of bookings to be accepted from all itinerary-fare classes in the network, consider the following mathematical formulation.

Consider the following notations:

ODF	Denotes an itinerary-fare class.
l	Denotes a flight in the network.
d_{ODF}	The demand for the itinerary-fare class *ODF*, which has a known statistical distribution.
f_{ODF}	The fare of the itinerary-fare class.
C_l	The seat capacity for flight *l*.
S_l	Denotes the set of all itinerary-fare classes that use the flight *l*.
x_{ODF}	The decision variable, which is the number of passengers to be accepted for the itinerary-fare class *ODF*.

The general mathematical formulation of the problem can be written as follows:

$$Maximize \quad E(\sum_{ODF} f_{ODF} \cdot \min\{x_{ODF}, d_{ODF}\}) \tag{14.1}$$

Subject to:

$$\sum_{ODF \in S_l} x_{ODF} \leq C_l \quad \forall\ l \tag{14.2}$$

$$x_{ODF} \geq 0 \quad \forall\ ODF \tag{14.3}$$

$$x_{ODF} \ \text{int} \quad \forall\ ODF \tag{14.4}$$

The objective function given in equation (14.1) is to maximize the total expected revenue from all bookings of the different itinerary-fare classes. The revenue is calculated by multiplying the fare of the itinerary-fare class by the number of bookings to be accepted for this booking class, given that the accepted number of bookings should not exceed the demand of the itinerary-fare class. The constraints

given in equation (14.2) ensure that the number of accepted bookings on the flight does not exceed the seat capacity of this flight. The constraints given in equations (14.3) and (14.4) ensure that the decision variables are positive integer numbers.

The objective function of this mathematical formulation depends on the demand distribution that is neither linear nor continuous. Accordingly, solving this problem is mathematically intractable. This mathematical formulation is simplified by only considering the expected value of the demand for the itinerary-fare class D_{ODF} instead of considering the demand distribution of the demand. In this case, the demand is considered to be deterministic. When the demand is deterministic, the problem can be formulated as a deterministic integer program with a linear objective function, as given below.

$$Maximize \quad \sum_{ODF} f_{ODF} \cdot x_{ODF} \tag{14.5}$$

Subject to:

$$\sum_{ODF \in S_l} x_{ODF} \leq C_l \qquad \forall \ l \tag{14.6}$$

$$x_{ODF} \leq D_{ODF} \qquad x_{ODF} \geq 0 \tag{14.7}$$

$$x_{ODF} \geq 0 \qquad \forall \ \ ODF \tag{14.8}$$

$$x_{ODF} \quad \text{int} \quad x_{ODF} \geq 0 \tag{14.9}$$

The objective function given in equation (14.5) is to maximize the total revenue from all bookings of the different itinerary-fare classes. The revenue is calculated by multiplying the fare of the itinerary-fare class by the number of bookings to be accepted for this booking class. The constraints given in equation (14.6), which are known as the capacity constraints, ensure that the number of accepted bookings on the flight does not exceed the seat capacity of this flight. The constraints given in equation (14.7), which are known as the demand constraints, ensure that the number of accepted bookings for any itinerary-fare class is less than the demand for this itinerary-fare class. The constraints given in equations (14.8)–(14.9) ensure that the decision variables are positive integer numbers. It is a common practice to solve the linear relaxation of the model by overlooking constraints in equation (14.9) rather than solving the integer program, which is usually hard to solve for large problems. The solution to this mathematical program gives the optimal number of passengers to be accepted for each itinerary-fare class.

An optimal solution to a linear program is a feasible solution with the largest objective function value (for a maximization problem). The value of the objective function for the optimal solution is said to be the value of the linear program. A linear program may have multiple optimal solutions, but only one optimal solution value. At the optimal solution, a constraint is defined to be a binding constraint, if

the left-hand side of the constraints is equal to its right-hand side. For instance, if a constraint represents a resource constraint, a binding resource constraint means that all the items of resource are used to maximize (for a maximization problem) the value of the objective function. In the deterministic linear program given above, a flight capacity constraint is binding at the optimal solution when all the seats of the flight are filled by passengers. A non-binding flight capacity constraint means that at the optimal solution, there are a few empty seats on the flight. Similarly, a demand constraint of any itinerary-fare class is binding when all booking requests for this itinerary-fare class are accepted. A non-binding demand constraint for any itinerary-fare class means that there are a few booking requests for this itinerary-fare class that are not accommodated.

At the optimal solution of the deterministic linear program, each binding constraint in the mathematical program has an associated dual price (also known as a shadow price, opportunity cost, or marginal cost). The dual price of the constraint gives an incremental change of the value of the objective function, when the right-hand side of the constraint is incremented by one. The dual price of any non-binding constraint is zero. For example, the dual price of the binding capacity constraint of any flight presented in equation (12.6) gives an incremental increase in the value of the total revenue when the number of seats of this flight is increased by one. Similarly, the dual price of the binding demand constraint of any itinerary-fare class presented in equation (12.7) gives an incremental increase in the value of the total revenue when the number of demand for this itinerary-fare class is increased by one.

Nesting with Network-based Models

A booking control policy based on the deterministic linear program can be constructed by setting booking limits for each Origin-Destination Fare (ODF) equal to the optimal number of passengers accepted for each itinerary-fare class. However, the booking limits on each flight have to be nested to consider the random nature of the demand. As mentioned earlier, the nesting of the different booking limits responds to an unexpected excess demand for one or more itinerary-fare classes. A nesting structure of booking classes is considered to always accept a high-value demand when there is space available on the flight. The booking limits of the different itinerary-fare classes are nested based on the contribution of each itinerary-fare class to the total network revenue. The procedure of determining the nesting mechanism of the different itinerary-fare classes is given below.

Williamson (1992)

Williamson suggests that nesting the different itinerary-fare combinations on each fight is made based on the incremental revenue that is generated if an additional seat is made available for each itinerary-fare class while everything else remains

unchanged. This incremental revenue is approximated to the incremental revenue obtained from increasing the mean demand of the itinerary-fare class by one. This incremental revenue is obtained by the dual price of the corresponding demand constraint (equation (12.7)) of this itinerary-fare class, which is given within the solution of the deterministic linear program given above.

Once the incremental revenue of each itinerary-fare class is determined, the itinerary-fare classes on each flight are nested according to their incremental revenue to the overall network. Then, a nesting booking control policy can be constructed for each itinerary-fare class as follows:

Let H_{ODF}^{l} be the set of itinerary-fare classes that have a higher rank than the itinerary-fate class *ODF* on flight *l*. Then, the nested booking limit, b_{ODF}^{l} for this itinerary-fare class cab be given as:

$$b_{ODF}^{l} = C_l - \sum_{ODF^{`} \in H_{ODF}^{l}} x_{ODF^{`}} \tag{14.10}$$

According to this nested booking limit, the bookings of an itinerary-fare class *ODF* can make use of all seats on flight *l*, except for the seats that are reserved for the higher-ranked itinerary-fare classes. When there is a booking request for any itinerary-fare class, all the booking limits of this itinerary-fare class are verified for each flight in this itinerary. The booking of this itinerary-fare class is accepted only if there is availability on all the flights of the itinerary.

De Boer (1999) and De Boer et al. (2002)

Similar to the previous method, the network contribution of the itinerary-fare classes are used to rank the classes in the nested booking structure. De Boer et al. use a different approach to estimate the network contribution. To estimate the network contribution of an itinerary-fare class, the following procedure is considered:

First, the dual prices of the capacity constraints of the flights that are included in this itinerary are obtained. As mentioned earlier, the dual price of the capacity constraint of a flight gives an incremental increase in the value of the total revenue when the number of seats of this flight is increased by one. Second, the sum of the dual prices of the flights that are included in the itinerary are computed, which gives the dual price of the itinerary, DP_{ODF}. The dual price of the itinerary gives the decrease in the total revenue when the seat capacity of this itinerary is decreased by one seat. It also gives the least contribution of a passenger on this itinerary. Third, the net contribution of the itinerary-fare class, $\overline{f}_{ODF}$, can be calculated as follows:

$$\overline{f}_{ODF} = f_{ODF} - DP_{ODF} \tag{14.11}$$

This equation means that the incremental revenue of a booking in any itinerary-fare class is calculated by replacing the least-valued passenger on this itinerary by a passenger from this itinerary-fare class. Once the nesting structure is determined, the booking control policy is determined as given above.

Illustrative Example

In this section, an example is given to illustrate the air carrier seat inventory control at the network level. For this purpose, consider the hypothetical air carrier network shown in Figure 14.2. The air carrier serves one hub station, M, and three spoke stations, A, B, and C, through six different flights. Each of these flights either originates or terminates at a hub station M, with seat capacity of 170 seats. Passengers traveling to and from the hub station are assumed to only consider the nonstop flight (itinerary) to and from the hub. Passengers traveling between two spoke stations are assumed to consider the single-stop itinerary that connects at the hub station M. It is assumed that each itinerary has two fare classes defined as the full fare (Y) and the discounted fare (T). Table 14.2 gives the list of all OD combinations served by this network. It also gives the different offered itinerary-fare classes. The fare value and the demand for each itinerary-fare class are also given. As shown in Table 14.2, there are 24 different itinerary-fare classes in the network. The objective is to determine the seat inventory control policy on each flight.

First, we solve the linear relaxation of the seat inventory control mathematical program [equations (14.5)–(14.8)]. We illustrate the problem formulation using the Excel optimization solver as shown in Table 14.3.

Step 1:
Consider the decision variable X for each ODF as given in column F. The decision variable X represents the number of passengers to be accepted for the itinerary-fare class *ODF*. Any initial values can be assumed for the decision variable. As shown in Table 14.3, we initialize the variable to be 15 passengers for each itinerary-fare class.

Step 2:
Next, we calculate the revenue from each itinerary-fare class by multiplying the number of passengers for each itinerary-fare class (column F) by the fare of the itinerary-fare class (column D). The result is given in column G.

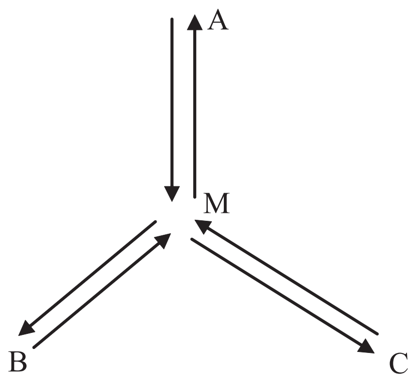


Figure 14. 2 A hypothetical air carrier network

Table 14.2 The list of all origin-destination combinations served by the network

Origin-Destination	Fare class	ODF	Fare ($)	Demand
A-C	Full	A-C Y	518	70
C-A	Full	C-A Y	522	66
M-C	Full	M-C Y	407	50
M-A	Full	M-A Y	411	50
C-M	Full	C-M Y	405	49
A-M	Full	A-M Y	412	52
B-C	Full	B-C Y	509	17
B-A	Full	B-A Y	510	16
C-B	Full	C-B Y	505	15
A-B	Full	A-B Y	517	15
B-M	Full	B-M Y	417	12
M-B	Full	M-B Y	423	10
A-C	Discounted	A-C T	311	32
C-A	Discounted	C-A T	309	35
M-C	Discounted	M-C T	252	100
M-A	Discounted	M-A T	259	100
A-M	Discounted	A-M T	254	97
C-M	Discounted	C-M T	251	104
B-C	Discounted	B-C T	300	33
B-A	Discounted	B-A T	307	33
A-B	Discounted	A-B T	306	31
C-B	Discounted	C-B T	304	30
M-B	Discounted	M-B T	255	20
B-M	Discounted	B-M T	257	22

Table 14.3 Problem formulation using Excel solver

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
	OD	**Fare class**	**ODF**	**Fare ($)**	**Demand**	**X**	**Revenue**	**flight MA**	**flight MB**	**flight MC**	**flight AM**	**flight BM**	**flight CM**	**flight MA**	**flight MB**	**flight MC**	**flight AM**	**flight BM**	**flight CM**
1	M-C	Full	M-C Y	407	50	15	6105			1				0	0	15	0	0	0
2	M-B	Full	M-B Y	423	10	15	6345		1					0	15	0	0	0	0
3	M-A	Full	M-A Y	411	50	15	6165	1						15	0	0	0	0	0
4	C-M	Full	C-M Y	405	49	15	6075						1	0	0	0	0	0	15
5	C-B	Full	C-B Y	505	15	15	7575		1				1	0	15	0	0	0	15
6	C-A	Full	C-A Y	522	66	15	7830	1					1	15	0	0	0	0	15
7	B-M	Full	B-M Y	417	12	15	6255					1		0	0	0	0	15	0
8	B-C	Full	B-C Y	509	17	15	7635			1		1		0	0	15	0	15	0
9	B-A	Full	B-A Y	510	16	15	7650	1				1		15	0	0	0	15	0
10	A-M	Full	A-M Y	412	52	15	6180				1			0	0	0	15	0	0
11	A-C	Full	A-C Y	518	70	15	7770			1	1			0	0	15	15	0	0
12	A-B	Full	A-B Y	517	15	15	7755		1		1			0	15	0	15	0	0
13	M-C	Discounted	M-C T	252	100	15	3780			1				0	0	15	0	0	0
14	M-B	Discounted	M-B T	255	20	15	3825		1					0	15	0	0	0	0
15	M-A	Discounted	M-A T	259	100	15	3885	1						15	0	0	0	0	0
16	C-M	Discounted	C-M T	251	104	15	3765						1	0	0	0	0	0	15
17	C-B	Discounted	C-B T	304	30	15	4560		1				1	0	15	0	0	0	15
18	C-A	Discounted	C-A T	309	35	15	4635	1					1	15	0	0	0	0	15
19	B-M	Discounted	B-M T	257	22	15	3855					1		0	0	0	0	15	0
20	B-C	Discounted	B-C T	300	33	15	4500			1		1		0	0	15	0	15	0
21	B-A	Discounted	B-A T	307	33	15	4605	1				1		15	0	0	0	15	0
22	A-M	Discounted	A-M T	254	97	15	3810				1			0	0	0	15	0	0
23	A-C	Discounted	A-C T	311	32	15	4665			1	1			0	0	15	15	0	0
24	A-B	Discounted	A-B T	306	31	15	4590		1		1			0	15	0	15	0	0
							133815	170	170	170	170	170	170	90	90	90	90	90	90

Step 3:
Calculate the total revenue, which is the sum of all the cells in column G, which gives the value of the objective function to be maximized by the solver.

Step 4:
The columns H through M give the flight-itinerary incidence matrix. Each column corresponds to one of the six flights in the network. A cell in a column is equal to one, if the flight is included in the corresponding itinerary. For example, the value of cell I12 is equal to one, because the flight MB is used within the itinerary connecting A to B (A-M-B).

Step 5:
The columns N–S give the number of passengers from each itinerary-fare class in which this flight is included. Each column corresponds to one of the six flights in the network. The number of passengers from an itinerary-fare class on a flight is calculated by multiplying the corresponding cell in the incidence matrix by the number of passengers to be accepted for the itinerary-fare class *ODF*. For example, column N is calculated by multiplying column H by column F.

Step 6:
Calculate the total number of passengers on each flight. For example, the total number of passengers on flight MA is calculated by the summation of all the cells in column N, which gives 90 passengers as shown in Table 14.3.

Step 7:
The mathematical program is defined to the solver as follows:

Maximize the total revenue which is given in cell G25 by changing the cells F1–F24, which correspond to the decision variables.

The constraints are defined as follows:

The flight capacity constraints are given by restricting the number of passengers allowed on a flight to less than the seat capacity of the flight. The value of each cell from N25–S25 should be less than the seat capacity of the corresponding flight (170 seats).

The demand constraints are given by limiting the number of accepted passengers for each ODF to be less than the demand of the ODF. These constraints are represented by limiting the value in each cell in column F to less than its corresponding cell in column E.

Finally, we constrain the values of the cells in column F to be greater than zero.

The solution to this problem is given in Table 14.4. Column F gives the total number of passengers to be accepted from each itinerary-fare class with a total revenue of $247,962.

Table 14.4 Problem solution

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
	OD	**Fare class**	**ODF**	**Fare ($)**	**Demand**	**X**	**Revenue**	**flight MA**	**flight MB**	**flight MC**	**flight AM**	**flight BM**	**flight CM**	**flight MA**	**flight MB**	**flight MC**	**flight AM**	**flight BM**	**flight CM**
1	M-C	Discounted	M-C T	252	100	0	0			1				0	0	0	0	0	0
2	M-C	Full	M-C Y	407	50	50	20350			1				0	0	50	0	0	0
3	M-B	Discounted	M-B T	255	20	20	5100		1					0	20	0	0	0	0
4	M-B	Full	M-B Y	423	10	10	4230		1					0	10	0	0	0	0
5	M-A	Discounted	M-A T	259	100	5	1295	1						5	0	0	0	0	0
6	M-A	Full	M-A Y	411	50	50	20550	1						50	0	0	0	0	0
7	C-M	Discounted	C-M T	251	104	10	2510						1	0	0	0	0	0	10
8	C-M	Full	C-M Y	405	49	49	19845						1	0	0	0	0	0	49
9	C-B	Discounted	C-B T	304	30	30	9120		1				1	0	30	0	0	0	30
10	C-B	Full	C-B Y	505	15	15	7575		1				1	0	15	0	0	0	15
11	C-A	Discounted	C-A T	309	35	0	0	1					1	0	0	0	0	0	0
12	C-A	Full	C-A Y	522	66	66	34452	1					1	66	0	0	0	0	66
13	B-M	Discounted	B-M T	257	22	22	5654					1		0	0	0	0	22	0
14	B-M	Full	B-M Y	417	12	12	5004					1		0	0	0	0	12	0
15	B-C	Discounted	B-C T	300	33	33	9900			1		1		0	0	33	0	33	0
16	B-C	Full	B-C Y	509	17	17	8653			1		1		0	0	17	0	17	0
17	B-A	Discounted	B-A T	307	33	33	10131	1				1		33	0	0	0	33	0
18	B-A	Full	B-A Y	510	16	16	8160	1				1		16	0	0	0	16	0
19	A-M	Discounted	A-M T	254	97	2	508				1			0	0	0	2	0	0
20	A-M	Full	A-M Y	412	52	52	21424				1			0	0	0	52	0	0
21	A-C	Discounted	A-C T	311	32	0	0			1	1			0	0	0	0	0	0
22	A-C	Full	A-C Y	518	70	70	36260			1	1			0	0	70	70	0	0
23	A-B	Discounted	A-B T	306	31	31	9486		1		1			0	31	0	31	0	0
24	A-B	Full	A-B Y	517	15	15	7755		1		1			0	15	0	15	0	0
							247962	170	170	170	170	170	170	170	121	170	170	133	170

Step 8:
We obtain the dual values of the flight capacity constraints that can be obtained from the sensitivity report of the Excel solution, as shown in Table 14.5. As mentioned above, the dual value is greater than zero only for binding constraints. As given by the solution, we have only four binding capacity constraints for flights MA, MC, AM, and CM. Each of these flights has 170 passengers at the optimal solution, which is equal to its seat capacity.

Table 14.5 The dual values of the flight capacity constraints

Flight	**flight MA**	**flight MB**	**flight MC**	**flight AM**	**flight BM**	**flight CM**
Dual value ($)	259	0	264	254	0	251

Step 9:
We calculate the total duality of each itinerary-fare class by adding the dual values of the flights that are included in the itinerary. For example, the total duality for the itinerary AC (A-M-C) is calculated by adding the dual value of flight AM ($254) and the dual value of flight MC ($264), which gives a total value of $518. Table 14.6 gives the total duality of each itinerary for the network. We also calculate the difference between the fare of the itinerary-fare class and its fare. We then sort the itinerary-fare classes based on the difference between these two values. This ranking gives the relative importance of each itinerary-fare class to the network revenue.

Step 10:
The nested seat availability for each itinerary-fare class on each flight is calculated based on the results obtained in Tables 14.4 and 14.6. For example for flight MA, there are six different itinerary-fare classes that use this flight, as given in Table 14.4. These itinerary-fare classes include M-A Y, B-M-A Y, C-M-A Y, M-A T, BM-A T, and C-M-A T. The optimal number of passengers to be accepted from each itinerary-fare class is as follows:

M-A Y	50
B-M-A Y	16
C-M-A Y	66
M-A T	5
B-M-A T	33
C-M-A T	0

The list of the itinerary-fare classes sorted according their relative importance to the network revenue, as given by Table 14.6 is as follows:

B-M-A Y	16

Table 14.6 Calculations of itinerary contributions

ODF	Flight Duality ($)						Total Duality ($)	Fare ($)	fare-duaility differnce ($)
	flight MA	flight MB	flight MC	flight AM	flight BM	flight CM			
M-B Y	0	0	0	0	0	0	0	423	423
B-M Y	0	0	0	0	0	0	0	417	417
A-B Y	0	0	0	254	0	0	254	517	263
B-M T	0	0	0	0	0	0	0	257	257
M-B T	0	0	0	0	0	0	0	255	255
C-B Y	0	0	0	0	0	251	251	505	254
B-A Y	259	0	0	0	0	0	259	510	251
B-C Y	0	0	264	0	0	0	264	509	245
A-M Y	0	0	0	254	0	0	254	412	158
C-M Y	0	0	0	0	0	251	251	405	154
M-A Y	259	0	0	0	0	0	259	411	152
M-C Y	0	0	264	0	0	0	264	407	143
C-B T	0	0	0	0	0	251	251	304	53
A-B T	0	0	0	254	0	0	254	306	52
B-A T	259	0	0	0	0	0	259	307	48
B-C T	0	0	264	0	0	0	264	300	36
C-A Y	259	0	0	0	0	251	510	522	12
A-C Y	0	0	264	254	0	0	518	518	0
M-A T	259	0	0	0	0	0	259	259	0
C-M T	0	0	0	0	0	251	251	251	0
A-M T	0	0	0	254	0	0	254	254	0
M-C T	0	0	264	0	0	0	264	252	-12
C-A T	259	0	0	0	0	251	510	309	-201
A-C T	0	0	264	254	0	0	518	311	-207

M-A Y	50
B-M-A T	33
C-M-A Y	66
M-A T	5
C-M-A T	0

The seat availability of an itinerary-fare class is then calculated by subtracting the number of seats reserved for the higher itinerary-fare classes from the flight

seat capacity. Accordingly, the protection level for each itinerary-fare class on flight MA is given as follows:

B-M-A Y	170
M-A Y	154
B-M-A T	104
C-M-A Y	71
M-A T	5
C-M-A T	0

It should be noted that the nesting booking limit is a heuristic to convert a non-nested solution from the mathematical program into a nested booking control policy. Allowing nesting eliminates the necessity to reserve seats for the different itinerary-fare classes as given by the solution of the mathematical program that violates optimality. To obtain an optimal solution, both the allocation and nesting decisions should be integrated in one formulation, which is an area under research (for example, dynamic programming techniques).

Primary Contributions

Early work for RM systems that consider more than one flight includes Ladany and Bedi (1977) who develop an optimal decision rule for allocating seats to passengers seeking to fly on a flight with an intermediate stop. Buhr (1982) addresses this problem for the two-segment case with two fare levels for each itinerary. For the multi-fare multi-leg case, Wang (1983) uses a marginal analysis approach, while Jessop (1985) uses Lagrange multipliers.

Glover et al. (1982) present the first network formulation for the airline RM, which depends on the assumption of deterministic demand. Dror et al. (1988) propose a deterministic network minimum cost flow formulation for the same problem. Wong et al. (1993) develop a network formulation for the case of the single-fare class and multiple itineraries. Extension to approximate for the multiple booking class case is also presented. The main drawback of these mathematical formulations is that they do not produce the clustering schemes of the different ODF combinations, the nesting structure of the booking classes, and the booking limit for each nested class. Belobaba (1987), Smith and Penn (1988), Williamson (1988), Simpson (1989), Williamson (1992), Vinod (1990), and Curry (1990) outline techniques for clustering ODFs into single-leg booking classes to achieve an approximation to network control. The main idea is to assign the ODFs to booking classes on the basis of their contribution to the total network revenue. Several options are considered to cluster the ODFs into booking classes. These options include assignment by the total value of the ODF, assignment by the estimated leg value after prorating the revenue of the itinerary to its legs (based on leg distance), and most commonly, assignment by the estimated net revenue contribution of the

ODF after subtracting any possible displacement effects (displacement cost). This displacement cost is equivalent to the dual prices from a deterministic network linear program formulation.

Curry (1990) describes a mathematical programming and marginal analysis formulation for the network RM problem to generate the distinct bucket allocations for different OD pairs. This methodology is implemented in several RM systems. Smith and Penn (1988) and Simpson (1989) propose the bid price concept for network RM. Williamson (1992) provides an extensive study of bid prices in comparison with other methodologies. He concludes that simple deterministic approximation methods based on average demand often outperform more advanced probabilistic heuristics. van Ryzin and Vulcano (1998) and Talluri and van Ryzin (1999) present the theoretical properties of the bid-price controls general stochastic network models based on Markov decision processes. Several types of approximations are developed and discussed in van Ryzin and Talluri (2003). De Boer et al. (2002) investigate the trade-off between the computation time and the aggregation level of demand uncertainty with examples of a multi-leg flight and a single-hub network. Cooper and Homem de-Mello (2003) combine the Markov decision processes and mathematical programming approaches. Bertsimas and Popescu (2003) investigate the dynamic policies for allocating scarce inventory to stochastic demand for multiple fare classes in a network environment so as to maximize total expected revenues. They propose and analyze a new algorithm based on approximate dynamic programming, which uses adaptive, non-additive bid prices from a linear programming relaxation.

References

Belobaba, P. 1987. Airline Yield Management: An Overview of Seat Inventory Control. *Transportation Science*, 21, 63-73.

Bertsimas, D. and Popescu, I. 2003. Revenue Management in a Dynamic Network Environment. *Transportation Science*, 37, 257-277.

Buhr, J. 1982. Optimal Sales Limits for Two-Sector Flights. *AGIFORS Symposium Proc. 22*.

Cooper, W.L. and Homem de-Mello, T. 2003. *A Class of Hybrid Methods for Revenue Management*. Working Paper No. 03-015, Northwestern University, Evanston, Illinois 60208-3119, U.S.A.

Curry, R.E. 1990. Optimal Airline Seat Allocation for Fare Classes Nested by Origins and Destinations. *Transportation Science* 24, 193-204.

De Boer, S.V. 1999. *Mathematical Programming in Airline Seat Inventory Control*. Master's Thesis, Vrije Universiteit, Amsterdam, The Netherlands.

De Boer, S.V., Freling, R. and Piersma, N. 2002. Mathematical Programming for Network Revenue Management Revisited. *European Journal of Operational Research*, 137(1), 72-92.

Dror, M., Trudeau D.P., and Ladany, S.P. 1988. Network Models for Seat Allocation of Flights. *Transportation Research B*, 22(4), 239-250.

Glover, F., Glover, R., Lorenzo, J., and McMillan, C. 1982. The Passenger Mix Problem in the Scheduled Airlines. *Interfaces*, 12, 73-79.

Jessop, A. 1985. *Optimal Seat Allocation for Airline Planning*. Presented to the 5th International Symposium on Forecasting, Montreal, Canada, 1985.

Ladany, S.P. and Bedi, D.N. 1977. Dynamic Rules for Flights with an Intermediate Stop. *Omega*, 5, 721-730.

Simpson, R.W. 1989. *Using Network Flow Techniques to Find Shadow Prices for Market and Seat Inventory Control*. Memorandum M89-1, MIT Flight Transportation Laboratory, Cambridge, MA.

Smith, B.C. and Penn, C.W. 1988. Analysis of Alternative Origin—Destination Control Strategies. *AGIFORS Symposium Proc.* 28.

Talluri, K. and van Ryzin, G. 1998. An Analysis of Bid-price Controls for Network Revenue Management. *Management Science*, 44, 1577-1593.

Talluri, K. and van Ryzin. G. 1999. A Randomized Linear Programming Method for Computing Network Bid Prices. *Transportation Science*, 33, 207-216.

van Ryzin, G. and Vulcano, G. 2003. *Simulation-based Optimization of Virtual Nesting Controls for Network Revenue Management*. Working paper DRO-2003-01, Columbia University, Graduate School of Business.

Vinod, B. 1990. Reservation Inventory Control Techniques to Maximize Revenues, in *IATA Third International Airline Yield Management Conference Proc.*, International Air Transport Association, London, England.

Wang, K. 1983. Optimum Seat Allocation for Multi-Leg Flights with Multiple Fare Types. *AGIFORS Symposium Proc.* 23.

Williamson, E.L. 1988. *Comparison of Optimization Techniques for Origin-Destination Seat Inventory Control.* Master's thesis, Flight Transportation Laboratory, Massachusetts Institute of Technology, Cambridge, MA.

Williamson, E.L. 1992. *Airline Network Seat Inventory Control: Methodologies and Revenue Impacts.* Ph.D. thesis, Flight Transportation Laboratory, Massachusetts Institute of Technology, Cambridge, MA.

Wong, J.T., Koppelman, F.S., and Daskin, M.S. 1993. Flexible Assignment Approach to Itinerary Seat Allocation. *Transportation Research*, 27B, 33-48.

Chapter 15
Ticket Distribution

Types of Ticket Distribution Channels

Once the air carrier determines the availability and pricing of the different itinerary-fare classes of the different flights, a distribution channel is needed to link the product to the potential customers. Ticket distribution channels are the media through which air carriers share travel information, including available itineraries and prices, to their prospective customers. Generally, the ticket distribution channels are classified into direct and indirect. In direct distribution, the air carriers make direct contact with their customers for ticket sales, and no intermediaries are involved in the sales process. With indirect distribution, intermediaries are involved in the sales process between the air carriers and the final customers.

Intermediaries are generally classified into wholesalers and agents. Wholesalers buy tickets in bulk from the different air carriers and sell them to the customers. They gain reasonable discounts from the airline companies by using their buying power. Wholesalers can adopt retailers or sell directly to final customers to sell tickets. The main disadvantage of wholesalers is that they give the air carriers less control over their tickets during the booking horizon of the different flights. Once the tickets are sold to the wholesalers, the air carriers are less able to change the prices or seat availability to respond to market competition. However, selling bulk seats in advance to wholesalers reduces the risk of air carriers having unsold seats on their flights.

The wholesaler intermediaries are more common with charter air carriers and tour operators. Agents are more common in the air carrier industry and are paid a commission each time they sell tickets on behalf of the air carrier. Agents have direct access to seat availability and the prices of different flights of the air carriers. Incentives are also given to agents when sales pass a predefined target for the air carrier. Agents could be selling on behalf of several competing air carriers. They may be tempted to use their market leverage to work harder to sell the products of the air carriers that pay them a higher commission.

Distribution channels can functionally be classified into two main categories. The first represents the air carriers' representatives, who promote ticket sales for one air carrier only. The second represents the travel agents, who simultaneously promote ticket sales for more than one air carrier. The two channels provide several ways to communicate with customers including the Internet, call-in telephone lines, and walk-in sales offices. These distribution channels differ in their cost, market penetration, type of customers, ease of use, and amount of information presented to customers.

Computer Reservation Systems

The Computer Reservation Systems (CRS) were developed in the early 1960s. They extend the air carriers' internal reservation systems that automatically mark the air carriers' seat availability. The CRS that include Apollo for United Airlines, SABRE for American Airlines, DATAS for Delta, and BABS for British Airways are examples of internal reservation systems. By the early 1970s most major air carriers had automated their internal reservations processes. However, travel agents were still dealing with air carriers by means of telephone links to reservation offices to check flight availability and pricing. By the late 1970s, with industry deregulation and the appearance of jet aircraft, cheap fares lead to an increased demand and consequently air carrier expansions. Still, travel agents did not have an efficient system to serve their increasing customers quickly and easily. In 1974, American Airlines (AA) initiated a joint feasibility study to explore the prospects for a jointly owned CRS for travel agents. In 1975, a study group, which included other air carriers, deemed the system economically practical. However, United Airlines withdrew from the project and announced their intension to provide their travel agents with their own CRS (Apollo). This step was followed by American Airlines that provided SABRE to travel agents.

Because many passengers were traveling through more than one air carrier, travel agents asked that the CRS system include the flight information of more than one air carrier. Therefore, the CRS owned by one air carrier could host the information of other air carriers. Also, air carriers that own CRS participated in the reservation systems of other air carriers to be accessible by more travel agents. These systems were called the Global Distribution Systems (GDS). The GDS provided travel information for a large number of air carriers across the globe in one system. Typically, the air carriers subscribed to one or more GDS, which accessed real-time seat availability and ticket price information from the air carriers' internal computer reservation systems. This information was made available to travel agents and other subscribed customers, who could effectively reach more customers on behalf of the air carriers, as shown in Figure 15.1. The GDS charged air carriers a pre-specified fee for each confirmed booking through the GDS. Also, the GDS might charge a subscription fee to travel agents in return for access to the information provided by the system.

It was found that travel agents tend to book their clients from the first computer screen of information on the CRS. Over 50 percent of the time, the bookings were made for the flight at the top of the first screen. The CRS owner used this fact and made sure that their flights were displayed more prominently on the first screen. Air carriers that own CRS could obtain a return on their investment through an increase in market share. It was argued that these CRS provided a solid way for the air carriers who owned them to increase their market share at the expense of their rivals. By the mid-1980s, the question of these early biased systems was becoming a controversial one. These complaints were raised by the air carriers that did not own reservation systems. The air carriers that owned systems argued

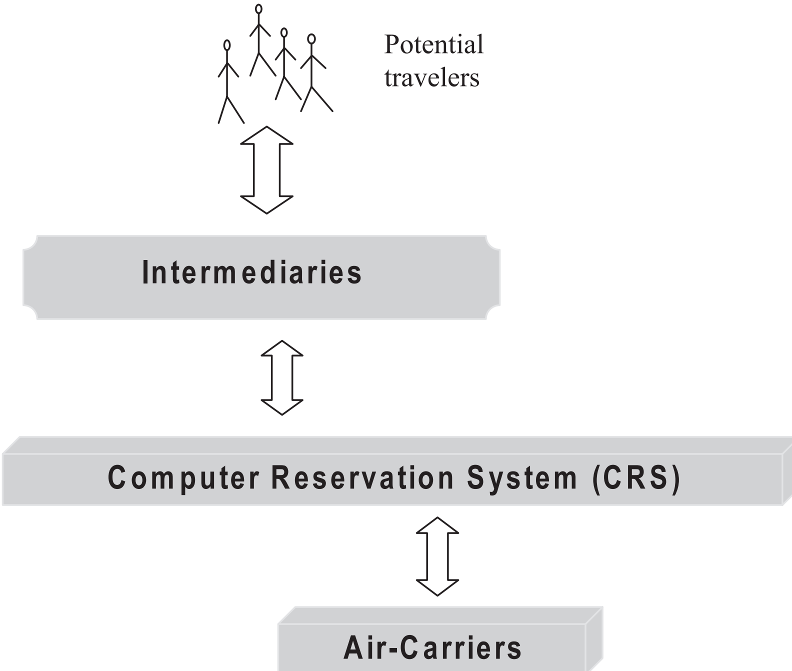


Figure 15.1 The role of distribution channels for the air carriers

that the returns they were getting through biased displays is to compensate for the investment they had made and the risk they had taken. Although the industry was deregulated, the biased CRS became subject to regulation by the US government, followed by the European Commission.

Once the display was neutralized by the regulation, the CRS-owned carriers were asking for another way to gain compensation for their investment costs. The system owners were authorized to charge other air carriers booking fees for every booking made through the reservation system they owned. After introducing the booking fees, the GDS were transformed into a highly profitable business. For example, the booking fees allowed American Airlines to invest in SABRE. At that time, more than 40 percent of the travel agents in the US. were using SABRE. This gave SABRE considerable power in the market. Other air carriers were concerned that the American Airlines' dominance through SABRE in the GDS industry would allow them to charge higher booking fees. In response to this threat, other carriers in the US and Europe merged their systems together to compete with SABRE. GALILEO, AMADEUS (European air carriers with several US-based air carriers) and WORLDSPAN (a smaller player set up by three US air carriers: Delta, Northwest, and the late TWA) are examples of these GDS. Currently, it is estimated that SABRE, GALILEO, and WORLDSPAN cover about 92 percent of the US markets.

The competition between the GDS still depends mainly on market shares, which depend on the number of agents that use the system. The GDS vendors tend to provide the service for free to agents who attract more customers. More incentives are provided by the GDS vendors to agents, and the air carriers are paying the price through increased booking fees. It is estimated that the incentives of the GDS to travel agents increased by 500 percent, and the value of the booking fee increased by about 31 percent from 1996 to 2002. In 2002, the major air carriers in the US posted net operating losses of $10 billion, and they paid $7 billion in booking fees.

The most recent development in the GDS industry can be summarized as follows. First, changes in the ownership of the GDS in recent years are because most air carriers that owned the GDS have wanted to maintain their liquidity, and have sold their shares in GDS. Second, air carriers have created and provided incentives to expand the use of various types of Internet-based applications that can bypass the GDS and their associated booking fees. These include the air carriers' own websites (air carrier.com), which bypass the GDS by using the air carriers' own internal reservation systems, and Orbitz, a travel technology company developed by a consortium of large US air carriers, has recently developed technology that allows it to book tickets without using a GDS. Meanwhile, several Internet-based travel agencies were developed. An example of these agencies includes Expedia (connected through Worldspan), Priceline (also known as Opaque for distressed inventory), and Travelocity—a subsidiary of one of the global distribution systems (SABRE). These Internet-based agencies use the GDS to book tickets, however they charge air carriers less than traditional travel agencies. The air carriers have encouraged some passengers to book a growing portion of tickets on less costly Internet-based booking sites, through various incentives. Accordingly, the industry has seen a considerable reduction of the proportion of the bookings that are coming through the traditional travel agency and GDS channel because of the Internet. Third, in another effort to cut distribution costs and in response to the increased share of Internet-based distribution channels, most air carriers in the US cut their sales commissions to traditional travel agents to reduce their distribution costs.

Current Practice in Ticket Distribution

The diagram given in Figure 15.2 presents the current practice of the air carriers' ticket distribution. As shown in the figure, there are four main entities that are included in the process. The diagram also highlights the interaction between these four entities by presenting the cost that is involved in the process. First, the consumer or prospective travelers seek their travel tickets through one of the available sources, which are highlighted as the travel agent. There are two main categories of travel agents that provide service to consumers. The first category includes the air carrier's own website (air carrier.com) and Orbitz, which directly accesses the internal reservation systems of the air carriers and bypasses the GDS.

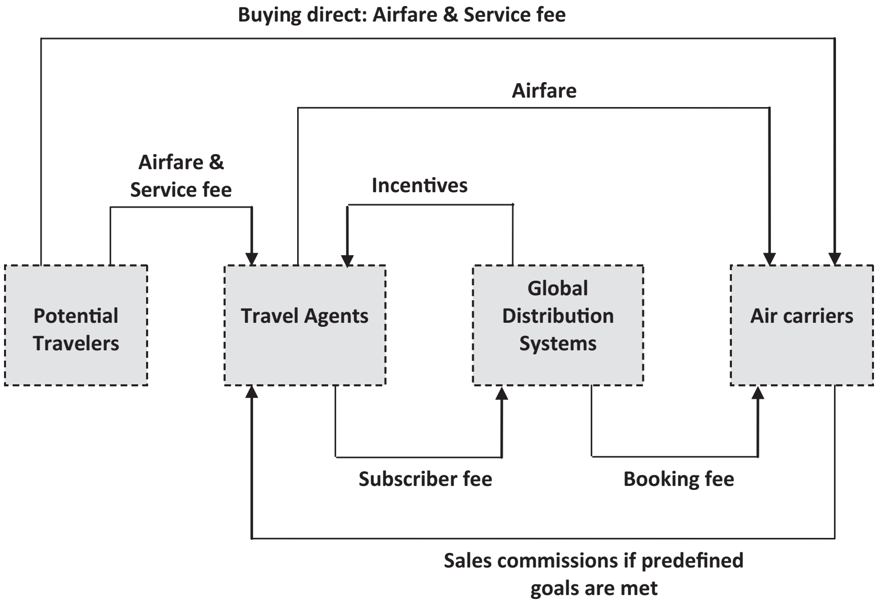


Figure 15.2 The current practice of the air carriers' ticket distribution

The other category includes the on-line travel agents and traditional travel agents that access travel information through the GDS. Consumers are expected to pay the required airfare when they book their tickets. In addition, most travel agents charge consumers a pre-defined service fee. The GDS are expected to charge subscription fees to those travel agents that need access to travel information. In many cases, to increase its market share of travel agents and sustain competition with other GDS, the travel agents are paid incentives by the GDS for their loyalty. Also, the GDS charge air carriers a booking fee for each booking performed to the air carrier through the GDS. Finally, air carriers might select to pay sales commissions to travel agents to boost their bookings.

Primary Contributions

Copeland and McKenney (1988) discuss the evolution of airline reservation systems. Durham (1995) outlines the future of ticket distribution channels. A recent study by GAO (2003) examines changes in the airline ticket distribution industry since the late 1990s. In particular, it addresses the effects on airlines, the impact of these changes on travel agents and consumers, and what the relationship between GDS booking fees and related costs suggest about the use of market power. An earlier study along the same line is given in GAO (1999).

The impact of ticket distribution on the airline RM of the airlines is given by Boyd and Bilegan (2003), who study the relationship between RM and e-commerce; and Boyd (2004), who addresses the impact of the changes in ticket distribution on pricing and RM models. Smith et al. (2001) discuss the impact of e-commerce and operations research on airline planning, marketing, and distribution. It is concluded that the availability of reliable, low-cost communications via the Internet is providing new modeling challenges within the airline industry.

Abdelghany and Abdelghany (2007) present a simulation model to evaluate the ticket distribution strategies of the airlines. The model examines the trade-offs between two common types of ticket distribution channels: 1) distribution channels with high market penetration and high competition among subscribed airlines (for example, Travelocity, Expedia, and Orbitz) versus 2) distribution channels with low market penetration and low airlines competition (for example, airlineName.com). Results show that both market penetration of the distribution channel and the level of competition between airlines within the distribution channel have considerable effect on the number of reservations of the airline. It is concluded that it is more beneficial for an airline to target distribution channels with high market penetration even if the level of competition within this channel is high.

References

Abdelghany, A., and Abdelghany, K. 2007. Evaluating Airlines Ticket Distribution Strategies: A Simulation-based Approach. *International Journal of Revenue Management*, 1(3), 231-246.

Boyd, A. 2004. Future of Revenue Management: Dramatic changes in distribution will require renewed focus on pricing and revenue management models. *Journal of Revenue and Pricing Management*, 3(1), 100-103.

Boyd, A. and Bilegan, I. 2003. Revenue Management and E-Commerce. *Management Science*, 49(10), 1363-1386.

Copeland, D., and McKenney, J. 1988. Airline Reservations Systems: Lessons from History, *MIS Quarterly*, 12(3), 353-370.

Durham, M. 1995. The Future of Sabre. In *Handbook of Airline Economics*, Aviation Week Group, McGraw-Hill, 485-491.

GAO, 1999. United States General Accounting Office. Domestic Aviation: Effects of Changes in How Airline Tickets Are Sold. GAO/RCED-99-221, U.S. General Accounting Office, 441 G Street NW, Room LM, Washington, D.C. 20548.

GAO, 2003. United States General Accounting Office. Airline Ticketing: Impact of Changes in the Airline Ticket Distribution Industry. Report to Congress 03-749, U.S. General Accounting Office, 441 G Street NW, Room LM, Washington, D.C. 20548.

Smith, B.C., Günther, D.P, Rao, B.V., and Ratlife, R.M. 2001. E-Commerce and Operations Research in Airline Planning, Marketing, and Distribution. *Interfaces*, 31(2), 37-55.

This page has been left blank intentionally

Chapter 16
Sales Contracts

Introduction

Over the last decade, business travel is estimated to account for about one-half of all US domestic air carriers' trips and two-thirds of the air carrier industry's passenger revenues (Bender and Stephenson 1998, Stephenson and Fox 1992). Several initiatives are implemented by commercial air carriers to attract corporate business travelers. A common initiative is to provide incentives to business travelers in the form of ticket discounts and cabin upgrades. A contract is signed between the air carrier and different corporations for this purpose. According to this contract, the air carrier provides a fare discount to the employees of the corporation when they select the air carrier for their travel. In return, the corporation has to guarantee that a portion of its travel is done through the air carrier. Such an agreement helps the corporations reduce their travel expenses and allows the air carrier to increase its travel share in the different markets.

On a periodic basis, an air carrier initiates and renews negotiated contracts with different corporations. In most cases, the contract with each corporation is evaluated independently from other contracts. The air carrier and the corporation perform several rounds of negotiation to determine the discount and the amount of travel share in each market, possibly at the cabin class level. The air carrier seeks to minimize the amount of discount given to the corporation and maximize its guaranteed travel share from the corporation's travel budget. The corporation seeks to obtain the best discount deal before it commits its travel to any air carrier.

Most major air carriers depend on ad hoc quantitative as well as some qualitative rules in evaluating their contracts with corporations to promote business travel. The air carriers usually evaluate contracts with corporations individually (that is, one corporation at a time), where two main steps are considered. The first step is to decide on the preliminary values for the ticket discounts to be offered to each corporation and the associated share of the corporation's business travel that the air carrier might obtain. The discount and travel share of each corporation could be decided either at the corporation level, the market level (that is, the OD), or the cabin class level. In most cases, the amount of discount and the associated travel share are determined after cycles of negotiation between the air carrier and the corporation. The air carrier determines the value of the offered discount based on the size of the corporation, the corporation's historical business travel pattern in the different markets, and the strength of competition with other air carriers in these markets. The corporation then decides on the expected discount after evaluating discount offers provided by all competing air carriers. Determining the

travel share is usually not trivial, as it depends on qualitative factors such as the significance of the travel policy of the corporation and its ability to convince its employees to travel by the contracted air carrier.

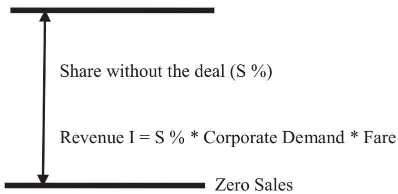
Factors Affecting Contract Evaluation

When the air carrier and the corporation decide on the preliminary discount and the associated travel share, the air carrier further analyzes these values to estimate the possible increase to its revenue. Three main factors are considered while evaluating any proposed contract. First, the dilution in revenue resulting from offering the discounted fares should be less than the extra revenue associated with attracting new demand. The increase in the air carrier's revenue associated with attracting new trips due to the offered discount should be greater than the loss resulting from the portion of this demand used to pay a full fare before signing the contract. Figure 16.1 illustrates the amount of revenue that an air carrier could achieve before and after signing a contract with a single corporation. In the case where no contract is signed, the revenue is calculated as the multiplication of the air carrier's travel share, the total corporations' travel demand, and the price of a full ticket. If a contract is signed with the corporation, an additional demand could be generated. The revenue in this case is calculated as the sum of the revenues generated from the original demand and the additional demand. Both demand types pay the discounted fare. In addition, the revenue loss associated with the possible displacement of the current passengers due to accepting new business passengers has to be considered. The incremental revenue is defined as the difference between the revenues generated in the two scenarios.

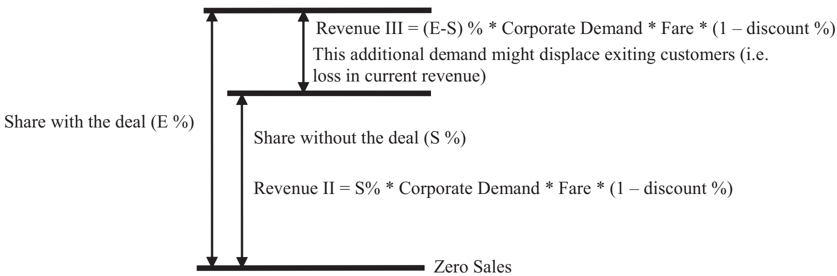
Second, because the attracted demand due to new contracts might displace the current leisure or business travelers, the new demand should be more profitable than the current demand. Also, the revenue of the displaced passengers should be less than the revenue from new business passengers associated with the new contracts. The next section presents the methodology used in calculating the displacement cost.

Finally, the revenue margin of any contract, which is defined as the total revenue from the contract minus the total cost divided by the total cost, should be greater than a predefined threshold. This revenue margin is useful for the air carrier as a measure to estimate the return rate of its investment from the corporation contract.

After this analytical evaluation is performed, if the deal is not profitable to the air carrier or it does not satisfy the revenue margin threshold, another round of negotiation might start with the corporation to improve the terms of the contract (for example, to reduce the offered discount or ask for a higher share). However, in the case that the corporation refuses to change the terms of the contract, the air carrier might select not to implement the contract with the corporation.



<u>Case A:</u> **Before the Contract**



<u>Case B:</u> **After the Contract**

Incremental Revenue = Revenue after the Contract – Revenue before the Contract
= (Revenue II + Revenue III – Displacement Cost) – Revenue I
= (E% - S%) * Corporate Demand * Fare * (1 – discount %) – S% * Corporate Demand * Fare * discount % - Displacement Cost

Figure 16.1 Illustration of incremental revenue calculation

Computation of the Displacement Cost

The methodology considered in contract evaluation depends on minimizing the displacement of the current leisure and business trips by the potential net demand associated with signing new corporation contracts. If the net generated demand is greater than zero, the new passengers are attracted to the air carriers. If no empty seats are available for these new business customers, they will displace the current leisure passengers that usually pay low discounted fares. Accordingly, the air carrier must ensure that the revenue from the new business customers is higher than that revenue generated by the displaced passengers, which necessitates an accurate computation of the revenue generated by the possibly displaced passengers.

If the displaced demand is known with certainty, the revenue can be estimated as the sum of the fares paid by the displaced passengers. However, the demand prediction is always uncertain, a fact that complicates estimating the number of

displaced seats and the associated revenue (Smith el al. 1992, McGill and Van Ryzin 1999).

In this section, we present the methodology used in estimating the revenue of the displaced passengers, which is also known as the displacement cost. The methodology is based on the idea that each seat in a flight has a probability of being filled by the predicted demand. The probability that a given seat *c* in a flight is filled is a function of the flight capacity and the expected demand of this flight. This probability is expected to increase as the demand increases and the flight capacity decreases. Accordingly, the expected marginal revenue generated from seat *c* in market *j* (defined as EMR_{cj}) could be calculated as the product of this probability and the full fare value of this seat.

However, it should be noted that the demand for a flight is not uniform, and it usually includes different classes of passengers that are willing to pay different fare values. Thus, EMR_{cj} is the highest revenue to be generated from all demand classes. This assumption complies with the way most air carriers' RM and inventory control systems work; they always reserve seats to meet the demand of high yield passengers (Smith et al. 1992, McGill and Van Ryzin 1999).

To explain further how EMR_{cj} is calculated, consider the following example in which a market is served by one flight with 40 seats. Assume that three different classes of passengers (A, B, and C) are expected to show up for this flight. Table 16.1 gives the average number of passengers expected for each type and the average fare they are willing to pay. Assume that the arrival of this demand is distributed according to the Poisson distribution (Walpole et al. 2002). Table 16.2 gives the probability that each seat will be filled by each of the three demand classes independently. These probabilities are calculated according to a Poisson distribution with the mean equal to the expected demand for each class. One should note that as the seat index in the reservation system increases, the probability decreases that this seat will be filled. The table also gives the expected revenue from each seat for each of the three demand classes. As mentioned earlier, the seat revenue is the product of the probability of filling the seat by a certain demand class and the average fare paid by this demand class. Once this value is calculated for each demand class, the expected seat revenue is determined as the highest revenue among all the demand classes, which is obtained by sorting all revenues from all demand classes and selecting the top 40 values. The top of the

Table 16.1 Expected demand and average fare for three hypothetical demand classes

Demand Class	Expected Demand (Passengers)	Average Fare ($)
A	20	200
B	12	300
C	10	500

Table 16.2 Calculation of expected seat revenue

Seat #	Probability of Filling Seat			Expected Seat Revenue ($)			EMR ($)	V ($)
	Class A	Class B	Class C	Class A	Class B	Class C		
1	1.00	1.00	0.98	200.00	300.00	490.84	490.84	8957.97
2	1.00	1.00	0.91	200.00	299.98	454.21	454.21	8467.12
3	1.00	1.00	0.76	200.00	299.84	380.95	380.95	8012.91
4	1.00	1.00	0.57	200.00	299.31	283.26	300.00	7631.97
5	1.00	0.99	0.37	200.00	297.72	185.58	299.98	7331.97
6	1.00	0.98	0.21	199.99	293.90	107.43	299.84	7031.99
7	1.00	0.95	0.11	199.95	286.25	55.34	299.31	6732.15
8	1.00	0.91	0.05	199.84	273.15	25.57	297.72	6432.84
9	1.00	0.84	0.02	199.58	253.49	10.68	293.90	6135.12
10	1.00	0.76	0.01	199.00	227.28	4.07	286.25	5841.22
11	0.99	0.65	0.00	197.84	195.83	1.42	283.26	5554.96
12	0.98	0.54	0.00	195.72	161.52	0.46	273.15	5271.70
13	0.96	0.42	0.00	192.20	127.21	0.14	253.49	4998.55
14	0.93	0.32	0.00	186.77	95.54	0.04	227.28	4745.06
15	0.90	0.23	0.00	179.03	68.39	0.01	200.00	4517.78
16	0.84	0.16	0.00	168.70	46.68	0.00	200.00	4317.78
17	0.78	0.10	0.00	155.79	30.39	0.00	200.00	4117.78
18	0.70	0.06	0.00	140.59	18.89	0.00	200.00	3917.78
19	0.62	0.04	0.00	123.72	11.22	0.00	200.00	3717.78
20	0.53	0.02	0.00	105.95	6.38	0.00	199.99	3517.78
21	0.44	0.01	0.00	88.18	3.48	0.00	199.95	3317.80
22	0.36	0.01	0.00	71.26	1.82	0.00	199.84	3117.85
23	0.28	0.00	0.00	55.88	0.91	0.00	199.58	2918.00
24	0.21	0.00	0.00	42.50	0.44	0.00	199.00	2718.42
25	0.16	0.00	0.00	31.35	0.21	0.00	197.84	2519.42
26	0.11	0.00	0.00	22.44	0.09	0.00	195.83	2321.58
27	0.08	0.00	0.00	15.58	0.04	0.00	195.72	2125.75
28	0.05	0.00	0.00	10.50	0.02	0.00	192.20	1930.03
29	0.03	0.00	0.00	6.87	0.01	0.00	186.77	1737.83

Table 16.2 *Concluded*

Seat #	Probability of Filling Seat			Expected Seat Revenue ($)			EMR ($)	V ($)
	Class A	Class B	Class C	Class A	Class B	Class C		
30	0.02	0.00	0.00	4.36	0.00	0.00	185.58	1551.06
31	0.01	0.00	0.00	2.69	0.00	0.00	179.03	1365.47
32	0.01	0.00	0.00	1.62	0.00	0.00	168.70	1186.45
33	0.00	0.00	0.00	0.95	0.00	0.00	161.52	1017.75
34	0.00	0.00	0.00	0.54	0.00	0.00	155.79	856.23
35	0.00	0.00	0.00	0.30	0.00	0.00	140.59	700.44
36	0.00	0.00	0.00	0.16	0.00	0.00	127.21	559.85
37	0.00	0.00	0.00	0.08	0.00	0.00	123.72	432.64
38	0.00	0.00	0.00	0.04	0.00	0.00	107.43	308.92
39	0.00	0.00	0.00	0.02	0.00	0.00	105.95	201.49
40	0.00	0.00	0.00	0.01	0.00	0.00	95.54	95.54

EMR column in Table 16.2 represents the highest value a passenger is willing to pay for a seat on this flight, while the bottom of the column represents the lowest-yield passenger.

After the expected seat revenue is calculated, the total displaced revenue V_j can be estimated by adding up the EMR_{cj} values for the seats that are to be filled by business passengers generated by the new contracts. It should be clear that the low-revenue passengers are the ones that are most likely to be displaced by business passengers generated by the new contracts. Accordingly, V_j is calculated by adding up the EMR_{cj} for the low-revenue seats for a number of seats that is equal to the number of business passengers generated by the new contracts T_j. For instance, in the previous example, if $T_j = 10$, V_j is calculated as the sum of EMR_{cj} for the last ten seats (31 to 40), which is equal to $1,365.

It should be noted that, the way EMR_{cj} is calculated implies that V_j is not linear with respect to the seat index c. To relax this non-linearity, several approaches can be used. For instance, a uniform value of the displacement cost per seat DC_j for all seats in market j is proposed. This value is calculated as the average EMR over all seats in market j, which is calculated as follows:

$$DC_j = (\sum_{1}^{N_j} EMR_{cj}) / N_j \quad \forall\, j \tag{16.1}$$

Where, N_j is the number of seats in market j. Other approaches that are used by the air carriers are to use only the average of the lowest 50 percent or 25 percent of the seats. Other approaches could be using a step-wise function, which categorizes

seats into buckets and assigns an average displacement cost for each bucket. Given the average displacement cost per seat in market j, the total displacement cost of T_j seats can then be computed as follows:

$$V_j = T_j \cdot DC_j \qquad \forall\ j \tag{16.2}$$

Primary Contributions

One main drawback of the current practice is that the current procedure evaluates the different contracts individually, and ignores the trade-off among the different proposed contracts and their aggregate impact on the profitability of each individual market. In addition, evaluating the contract in an ad hoc sequential procedure results in portfolios that are far from optimal and confines the potential revenues. For instance, an early signing of a contract could result in rejecting a more profitable contract evaluated later in time. Furthermore, the use of the margin threshold to decide on accepting and rejecting contracts could be misleading, as some profitable contracts might be rejected because they do not satisfy the threshold by a small margin. Finally, evaluating the contract for each corporation on an individual basis is a repetitive time-consuming procedure, one that could be efficiently performed by simultaneously evaluating the proposed pool of contracts.

Abdelghany et al. (2007) presented a model for managing an air carrier's portfolio of corporation contracts that overcomes these shortfalls of current practice. The model extends the current practice's procedure by evaluating the complete portfolio of contracts simultaneously. The output of the model represents the optimal contract portfolio to be signed by the air carrier. The problem is formulated similar to the traditional knapsack problem, where the revenue is maximized given the limited seat availability on each market. The model captures the trade-off between accepting new contracts and the possible displacement of current demand. It evaluates deals at different disaggregate levels, including the OD markets and cabin classes. The input to the model is the list of new and existing (close to expiring) contracts. Each contract is defined in terms of the negotiated amount of discount to be offered by the air carrier and the associated business travel share that the air carrier would receive from the corporation's travel inventory. The model output is an accept-reject decision for each contract in the input list.

References

Abdelghany, A. and Abdelghany, K. 2007. Modeling Air Carrier's Portfolio of Business Travel. *Journal of Revenue and Pricing Management*, 6(1), 51-63.

Bender, A. and Stephenson, F. 1998. Contemporary Issues Affecting the Demand for Business Air Travel in the United States. *Journal of Air Transport Management*, 4 (2), 99-109.

McGill, J.I. and Van Ryzin, G.J. 1999. Revenue Management: Research Overview and Prospects. *Transportation Science*, 33 (2), 233-256.

Smith, B.C., Leimkuhler, J.P., and Darrow, R.M. 1992. Yield Management at American Airlines. *Interfaces*, 22 (1), 8-31.

Stephenson, F. and Fox, R. 1992. Corporate Strategies for Frequent-flier Programs. *Transportation Journal*, 32 (1), 38-50.

Walpole, R.E., Myers, R.H., Myers, S.L., Ye, K., and Yee, K. 2002. *Probability and Statistics for Engineers and Scientists* (7th Edition), Prentice Hall, Upper Saddle River, NJ.

Chapter 17
Code-share Agreements

Introduction

Code share is a partnership agreement between two air carriers; the operating carrier and the marketing carrier. The operating carrier allows the marketing carrier to promote and sell seats on some of its flights (Ito and Lee 2007). Code share brings a benefit to both the marketing and operating carriers of the agreement. It allows the marketing carrier to increase network coverage by gaining access to the routes of the code-share partner. The operating carrier uses the agreement to increase its flights' load factor and provides an additional stream of revenue. In addition, code share provides relatively seamless travel for customers when their travel requires connecting between the flights of different air carriers. Figure 17.1 shows a network structure of a hypothetical operating carrier *o* that could have code-share agreement with a marketing carrier *m*. It is possible that the operating carrier participates simultaneously in multiple code-share agreements with different marketing carriers (for example, the alliance). Thus, several marketing carriers promote the flights of the operating carrier at the same time.

Types of code-share agreements

Vinod (2005) defined three different types of code-share agreements: the hard block, soft block, and free sale. In the hard block agreement, the marketing carrier purchases a predetermined number of seats at a predetermined price from the operating carrier. Then, the marketing carrier is free to sell those seats at whatever price it chooses (subject to regulatory issues). In the soft block agreement, the size of the block is not determined in advance and is expected to vary based on the operating carrier's anticipated demand level. Finally, in the free sale (also known as free flow) agreement, the operating carrier gives the marketing carrier access to its seat inventory, which the marketing carrier can promote under its own code. In some cases, the operating carrier uses a seat inventory control mechanism to block certain booking classes from being accessed by the marketing carrier.

Challenges of Code-share Agreements

Regardless of the type of code-share agreement, in order to gain the most out of code sharing, each participating air carrier should answer several questions,

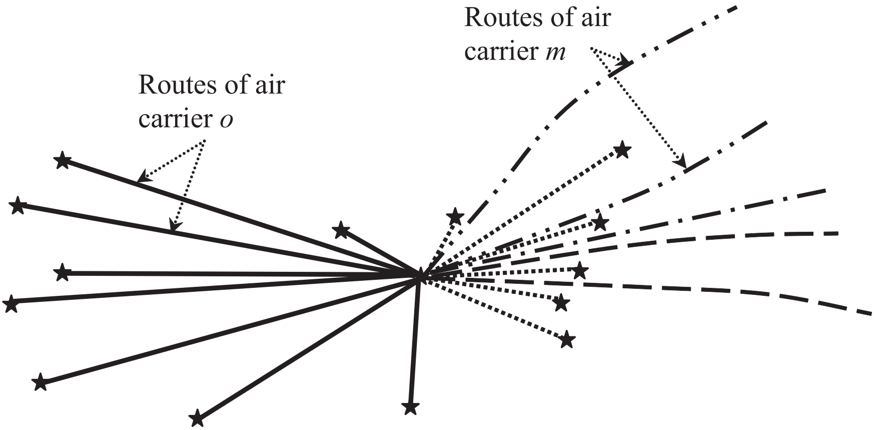


Figure 17.1 The routes of a hypothetical air carrier participating in multiple code-share agreements

including which air carriers would be most beneficial to enter into a code-share agreement with, which flights would be accessible to each code-share partner, how many seats would be for sale by the different partners on each flight, what fare price would be given to each partner on each flight, and how many in-return seats should be requested on the flights of the different partners and the price of these seats.

In the early implementation of code-share agreements, the conventional wisdom has been that the greater the extent of code share, in terms of flights, connections, and passengers, the more the revenue is to the air carrier (Sivakumar 2003). Most code-share agreements are implemented based on the assumption that the agreement will generate additional demand that will fill empty seats on the participating flights. However, in some cases, this additional code-share demand would displace non-code-share passengers on the operating carrier's flights. Accordingly, it is crucial to each airline to evaluate the trade-off between the incremental revenue from the code-share agreement and the revenue loss due to displacing non-code-share passengers. This problem becomes more convoluted when the air carrier particpates in multiple code-share agreements with several marketing air carriers. In this case, the operating carrier must find the optimal seat allocation among the different partners to maximize its revenue and to minimize the chances of displacing the non-code-share passengers.

Most air carriers adopt ad hoc quantitative rules to evaluate the profitability of their code-share agreements. In most cases, the terms of the agreement are determined after cycles of negotiation between the operating carrier and the marketing carriers, where each air carrier seeks to maximize its own expected revenues. One main drawback of the current practice is that most airlines overlook the possible revenue loss from displacing non-code-share passengers. In addition, the different code-share agreements are evaluated individually, ignoring the trade-

off among the different proposed agreements and its overall impact on profitability. Evaluating the code-share agreement in a sequential procedure could result in portfolios that are far from optimal and restrict the potential revenues.

Formal Problem Definition

As shown in Figure 17.1, consider a commercial air carrier *o* that operates *I* flights. Each flight $i \in I$ is defined in terms of its seat capacity c_i, the average fare f_i, and the average passenger demand d_i. Consider *M* other air carriers that are interested in entering into code-share agreements with air carrier *o*. Each air carrier $m \in M$ operates a number of flights J^m that connect to the network of air carrier *o*. The average fare g_j^m of each flight j $\in J^m$ is assumed to be known. As part of the agreement with air carrier *o*, each air carrier $m \in M$ is interested in accessing a predefined number of seats s_i^m on flight $i \in I$ at a fare value of $(1-q_i^m).f_i$, where q_i^m is a possible discount given to air carrier *m* on flight *i*. In return, air carrier *o* is assumed to sell seats on some of the flights of air carrier *m* that connect to flight $i \in I$. We define the set of flights $J_i^m \in J^m$ as the subset of flights that are to be accessed by air carrier *o* in the network of air carrier *m* when air carrier *m* gets access to flight $i \in I$ in the network of air carrier *o*. Air carrier *o* is assumed to access a predefined number of seats n_j^m on each flight $j \in J_i^m$ of air carrier *m* at a certain fare equal to $(1-p_j^m).g_j^m$, where p_j^m is a possible discount that is given by air carrier *m* to air carrier *o* on flight *j*. The values of q_i^m and p_j^m are determined based on prorate agreements between the operating carrier *o* and each marketing carrier $m \in M$. Assume that the proposed code-share agreements generate T_i new passengers on each flight $i \in I$. This newly generated demand displaces at most V_i of the existing non-code-share passengers with revenue V_i. The displaced revenue V_i (also called the displacement cost of flight $i \in I$) is calculated as a function of the average demand d_i and the fare f_i.

The goal of air carrier *o* is to decide which code-share agreements to participate in to maximize its revenue. To achieve this goal, the operating carrier should consider the trade-off between the revenue from the different code-share agreements and possible displacement of the non-code-share passengers on its flights. In addition, the operating carrier should consider the amount of generated revenue from selling seats on the different flights of each code-share partner $m \in M$. We define the binary decision as variable x_i^m, which is equal to 1 if air carrier *o* (the operating carrier) is to accept the code-share agreement with air carrier *m* (the marketing carrier) on flight *i*; otherwise, the binary decision variable is equal to 0.

Mathematical Formulation

The problem described above can be presented mathematically as follows.

Maximize:

$$R = \sum_{m\in M}\sum_{i\in I} x_i^m s_i^m (1-q_i^m) f_i \; + \sum_{m\in M}\sum_{j\in J_i^m} x_i^m n_j^m (1-p_j^m) g_j - \sum_i V_i \qquad (17.1)$$

Subject to:

$$\sum_{m\in M} x_i^m s_i^m \le c_i \qquad \forall i \in I \qquad (17.2)$$

$$V_i = \sum_{k=1}^{T_i} EMR_{ki} \qquad \forall \mathrm{i} \qquad (17.3)$$

$$T_i = \sum_{m\in M} x_i^m s_i^m \qquad \forall i \qquad (17.4)$$

$$x_i^m = x_l^m \qquad \forall\, l \ne i,\ m \qquad (17.5)$$

$$x_i^m \in \{0,1\} \qquad \forall\, i,\ m \qquad (17.6)$$

The objective function in equation (17.1) maximizes the operating carrier's additional revenue *R* that is associated with participating in code-share agreements with air carriers in *M*. The objective function is composed of three components. The first represents the revenue gained by selling seats to the code-share partners. The second is the revenue achieved by selling seats on the partners' flights. Finally, the third component represents the possible revenue loss due to displacing the non-code-share passengers. Constraints in equation (17.2) ensure that the newly generated demand on any flight $i \in I$ must be less than its available seat capacity. Constraints in equation (17.3) describe the displacement cost for each flight $i \in I$. Constraints in equation (17.4) are to estimate the additional demand due to the new code-share agreements. Constraints in equation (17.5) are considered only if an all-or-nothing agreement is proposed; that is, the marketing carrier accesses all requested flights of the operating air carrier. Otherwise, the agreement is not implemented. Overlooking these constraints implies that the operating carrier can accept a code-share agreement on some flights and prevent access on others. Constraints in equation (17.6) ensure the integrality of the decision variables x_i^m ($\forall\ i, m$).

The way the displacement cost is calculated (see the previous chapter) implies that the formulation presented above involves a non-linear objective function. Unless, the displacement cost is approximated to obtain a linear formulation, a heuristic approach is used to solve the problem.

Primary Contributions

Most airline code-share literature focuses on studying the impact of code share on market competition, fare pricing, and cost structure of air carriers. Examples of these studies include Oum et al. (1996), Park (1997), Park and Zhang (2000), Brueckner (2001 and 2003), Chua et al. (2005), Goh and Yong (2006), and Ito and Lee (2007). The literature that considers modeling code-share agreements is relatively limited. For example, Wen and Hsu (2006) present an interactive airline network design procedure to facilitate bargaining interactions necessitated by international code-share alliance agreements. Sivakumar (2003) presents a code-share optimizer model used by United Airlines. It represents a network optimization tool that considers the complex interaction between the demand, fares, market shares, and prorate agreements that recommend optimal code-share levels. O' Neal et al. (2007) present a code-share flight-profitability system that automates the code-share flight-selection process for Delta Airlines. The system chooses a set of code-share flights that maximizes the total system revenue for Delta while satisfying the rules that the alliance, the government, and the airlines' unions have set. The new system for choosing code-shared flights seeks to increase Delta's revenue by up to $50 million per year while reducing the planning cycle to several hours. Abdelghany et al. (2009) presents an optimization model to evaluate the profitability of air-carriers' code-share agreements. The input to the model represents the terms of the different code-share agreements proposed to the air carrier. The model output provides a recommendation of the set of agreements to be accepted. The model considers the trade-off between the incremental revenue from the code-share agreements and the potential loss of revenue from the displacement of the non-code-share passengers. Two main sets of experiments are considered to present the model functionality. The first set of experiments investigates the impact of the different terms of the proposed agreements on the expected profitability. Based on the results of the conducted experiments, it is concluded that it is critical for the air carrier to determine the flights and number of seats that can be opened for code share to avoid displacing its high-revenue demand by the discounted demand of the partnering air carriers. In the second set of experiments, the model is used to compare the evaluation methodology for the agreements: concurrent versus sequential. The results indicate that evaluating the agreements simultaneously provides the air carriers with a better opportunity to optimize their code-share portfolio. An improvement of up to 30 percent can be achieved compared to when the agreements are evaluated sequentially.

References

Abdelghany, A., Sattayalekha, W., and Abdelghany, K. 2009. On Airlines Code-Share Optimization: A Modeling Framework and Analysis. *International Journal of Revenue Management*, 3(3), pp. 307-303.

Brueckner, J. 2001. The Economics of International Codesharing: An Analysis of Airline Alliances. *International Journal of Industrial Organization*, 19(10), 1475-1498.

Brueckner, J. 2003. International Airfares in the Age of Alliances: The Affects to Codesharing and Antitrust Immunity. *Review of Economics Statistics*, 85(1), 105-118.

Chua, C.L., Kew, H., and Yong, J. 2005. Airline Code-share Alliances and Costs: Imposing Concavity on Translog Cost Function Estimation. *Review of Industrial Organization*, 26(4), 461-487.

Goh, M. and Yong, J. 2006. Impacts of Code-share Alliances on Airline Cost Structure: A Truncated Third-order Translong Estimation. *International Journal of Industrial Organization*, 24(4), 835-851.

Ito, H. and Lee, D. 2007. Domestic Codesharing, Alliances and Airfares in the U.S. Airline Industry. *Journal of Law and Economics*, 50 (2), 355-380.

O'Neal, J.W., Jacob, M.S., Farmer, A.K., and Martin, K.G. 2007. Development of a Codeshare Flight-Profitability System at Delta Air Lines. *Interfaces*, 37(5), 436-444.

Oum, T.H., Park, J.H., and Zhang, A. 1996. The Effects of Airline Codesharing Agreements on Firm Conduct and International Air Fares. *Journal of Transport Economics and Policy*. 30(2), 187-202.

Park, J.H. 1997. The Effects of Airline Alliances on Markets and Economic Welfare. *Transportation Research E*, 33(3), 198-205.

Park, J.H. and Zhang, A. 2000. An Empirical Analysis of Global Airline Alliances: Cases in the North Atlantic Markets. *Review of Industrial Organization*, 16(4), 367-384.

Sivakumar, R. 2003 *Codeshare Optimizer—Maximizing Codeshare Revenues*. AGIFORS Schedule and Strategic Planning Study Group Meeting, Toulouse, France.

Vinod, B. 2005. Alliance Revenue Management. *Journal of Revenue and Pricing Management*, 4 (1), 66-81.

Wen, Y.H. and Hsu, C.I. 2006. Interactive Multiobjective Programming in Airline Network Design for International Airline Code-share Alliance. *European Journal of Operational Research*, 174(1), 404-426.

SECTION IV
Irregular Operations Management

This page has been left blank intentionally

Chapter 18
Ground Delay Programs and Collaborative Decision Making

Ground Delay Program

Airline management invests all of its effort to develop profitable schedules that are satisfactorily competitive in the different markets. This effort entails collecting and processing a significant amount of information related to market demand, air carrier competition, airport capacity, fleet size, maintenance requirements, crew working rules, aviation regulations, and so on. During normal operation conditions, the airline is expected to operate its planned schedule smoothly with slight or no modifications. However, when adverse weather conditions are anticipated, several safety regulations are imposed by the aviation administration to guarantee safe operation. These safety regulations are usually in the form of a capacity reduction on runways at airports that have adverse weather conditions. This capacity reduction represents a decrease in the number of flights that can land at the airport. Accordingly, the airline schedule is expected to be disrupted due to the imposed capacity reduction.

In the US, for example, flights that are scheduled to arrive at airports with adverse weather conditions are issued what is known as a Ground Delay Program (GDP). A GDP, which is set by the FAA, delays the departure of flights at their origin stations until weather conditions allow safe landing at their destination stations. Figure 18.1 shows an example of a GDP issued for a single flight that is planned to land during a period of an adverse weather condition. This flight is asked to delay its arrival time and its departure time, so that it arrives at the destination when the weather improves. The idea behind the GDP is that it is safer to hold a flight on the ground at its origin than to let it takeoff and be subject to airborne and landing difficulties at its destination due to the adverse weather conditions. The GDP is a key tool to balance and meter the flight arrivals by all air carriers and the capacity at airports experiencing a significant reduction in capacity mainly due to adverse weather conditions. The GDP reduces the number of arrivals during the period of adverse weather conditions at the airport by increasing the time intervals between the flights of all air carriers. Under the GDP, fewer landing slots are allowed during this time period. If feasible, more landing slots are to be used in the following period of good weather conditions. Hopefully, the demand is not high during the period of good weather conditions so that vacant landing slots in this period can be used. Figure 18.2 shows an example of a GDP issued for the flights arriving at a hypothetical airport. As shown in the figure, each arrow represents a flight arriving at the airport at a certain arrival time. The shape of the arrow indicates the air carrier that operates the flight. There are 14

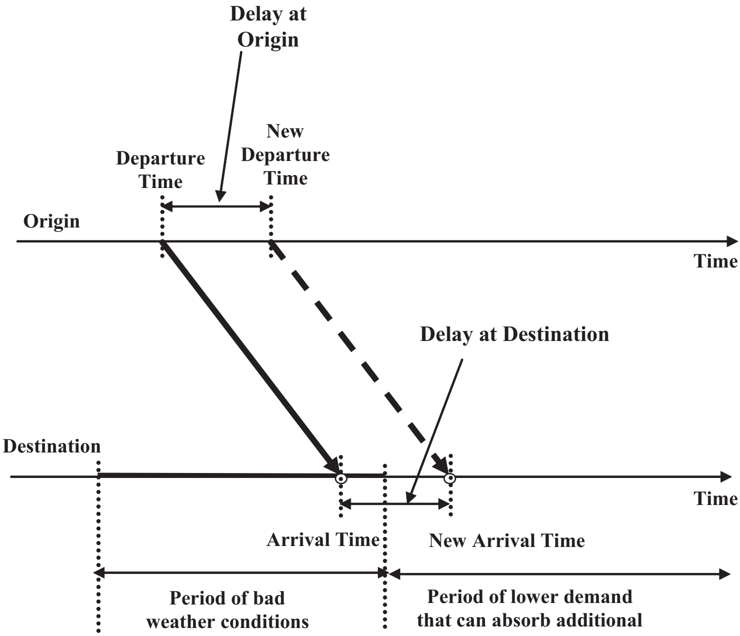


Figure 18. 1 An example of a ground delay program issued for a single flight

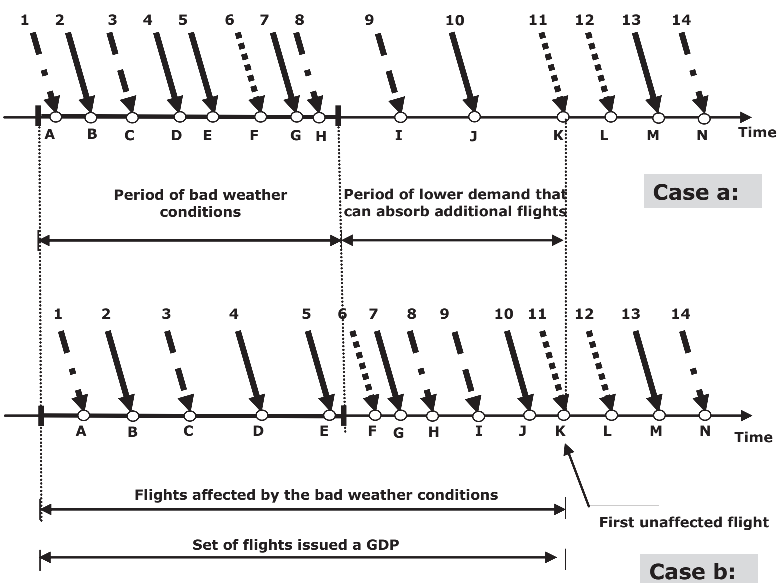


Figure 18. 2 An example of a GDP issued for the flights arriving at a hypothetical airport

flights that belong to five different air carriers. Case (a) of Figure 18.2 represents the originally scheduled arrival time of each flight. When adverse weather conditions are anticipated at the airport, the aviation administration issues a GDP to reduce the arrivals during this period. As shown in Case (b), only five arrivals are allowed during the period of adverse weather conditions. Also, additional landing slots are used later in the day when the weather conditions improve to accommodate more flights. According to this example, Flights 1 through 10 are issued a GDP. Therefore, these flights are asked to be delayed so they arrive at the airport at their newly scheduled time. Flights 11, 12, 13, and 14 are not affected by the GDP. Figure 18.3 shows another example of a GDP at a hypothetical airport, where the adverse weather conditions extend over a longer period. As shown in this figure, only ten landing slots can be used at the airport. Therefore, only ten flights can land at the airport and four flights need to be cancelled.

An important question is how the new slots are distributed among the different air carriers during the GDP. Initially, the distribution is based on the rule of first come, first served. Accordingly, air carriers provide the aviation administration with their planned (static) schedule of flight arrivals at the different airports. During a GDP, the available landing slots are allocated to the air carriers based on this static schedule. For example, in Figure 18.3, Case (a) shows the static schedule of flight arrivals, where 14 flights are scheduled to land at the airport. In

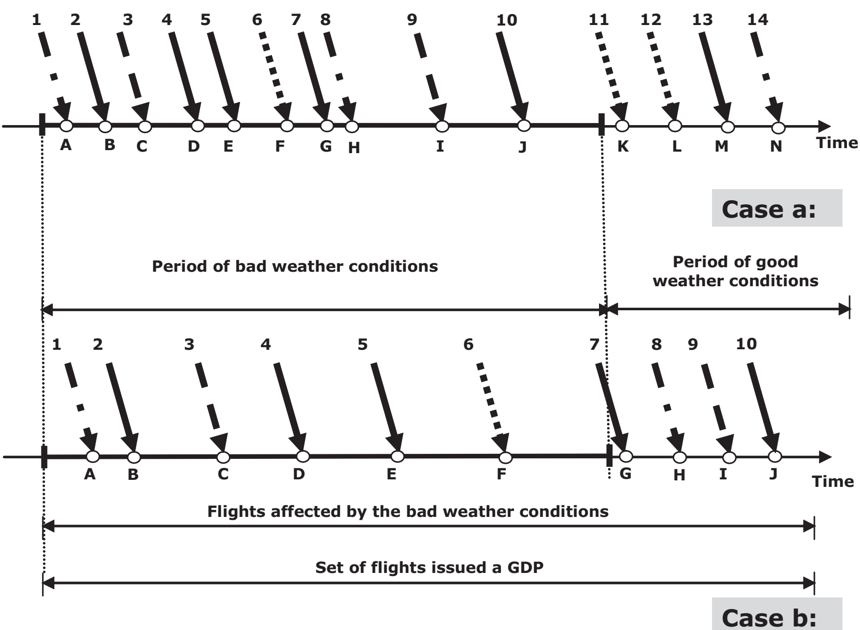


Figure 18. 3 Another example of a GDP at a hypothetical airport where the adverse weather conditions extend over a longer period

Case (b), during the GDP, ten landing slots are available at this airport (A through J). According to the first-come-first-served rule, flights 1 through 10 are assigned to slots A through J, respectively. No slots are available for flights 11 through 14. These flights have to be cancelled or diverted to other adjacent airports.

This approach has one major limitation. An air carrier might end up not using a slot that is allocated to it. For example, the air carrier might cancel the flight that is allocated to this slot due to an aircraft mechanical problem or lack of other resources. Not using this landing slot is a waste, because no other air carrier can benefit from it. Other air carriers can benefit from this slot only if they know about this cancellation ahead of time, so that they can plan their schedule accordingly. It is clear from this example that air carriers should provide the aviation administration with their updated (latest) schedules. Any information about flight cancellations or delays should be provided. Indeed, air carriers find there are disincentives for sharing their updated schedule information with the aviation administration.

First, air carriers know that the canceled flights are not allocated slots. It is in the air carrier's best interest not to report a flight cancellation but wait instead until the GDP is issued and a slot is given to the flight. Then, the air carrier cancels the flight and substitutes another flight into the vacant slot. Figure 18.4 explains this plan for a hypothetical air carrier. The air carrier that operates flight 5 has a plan to cancel this flight. In Case (b), if the air carrier reports to the aviation authority the cancellation of flight 5, flight 5 does not get a slot and flight 6, which is operated by another air carrier, is assigned to slot E. At the same time, flight 7, which is operated by the same air carrier that operates flight 5, is assigned to slot F. However, if this air carrier does not report the cancellation of flight 5, this air carrier is awarded slot E. Then, flight 7 can be moved up to slot E and saves the delay of flight 7, as shown in Case (c) of Figure 18.4. Second, air carriers know that the flights are allocated to slots based on the latest estimated arrival time of the flights. If the air carrier has a plan to delay a flight due to any reason such as an aircraft mechanical problem, it is better for the air carrier not to report this delay but wait instead until the GDP is issued then compare the proposed delay to the delay from the GDP. If the air carrier reports a proposed delay for one of its flights, this flight is given greater delay or doubly penalized due to the GDP. Figure 18.5 explains the difference between the scenarios of reporting a flight delay and not reporting a flight delay for a hypothetical air carrier. As shown in Figure 18.5, for Cases (a) and (b), the air carrier that operates flight 5 reports the proposed delay of flight 5. When the first-come-first-served rule is applied, flight 5 is assigned to slot F, as shown in Case (b). If this air carrier does not report the proposed delay of flight 5, flight 5 is assigned to slot E, which is scheduled anyway after the proposed delay of flight 5. However slot E is still better than slot F.

More importantly, by not sharing the real-time schedule information between the aviation administration and the air carriers, the aviation administration has limited capabilities to monitor or adjust GDPs once they are initially issued. The weather may improve or deteriorate all through the day, so having the updated schedule information would help to better manage flight arrivals in this constrained

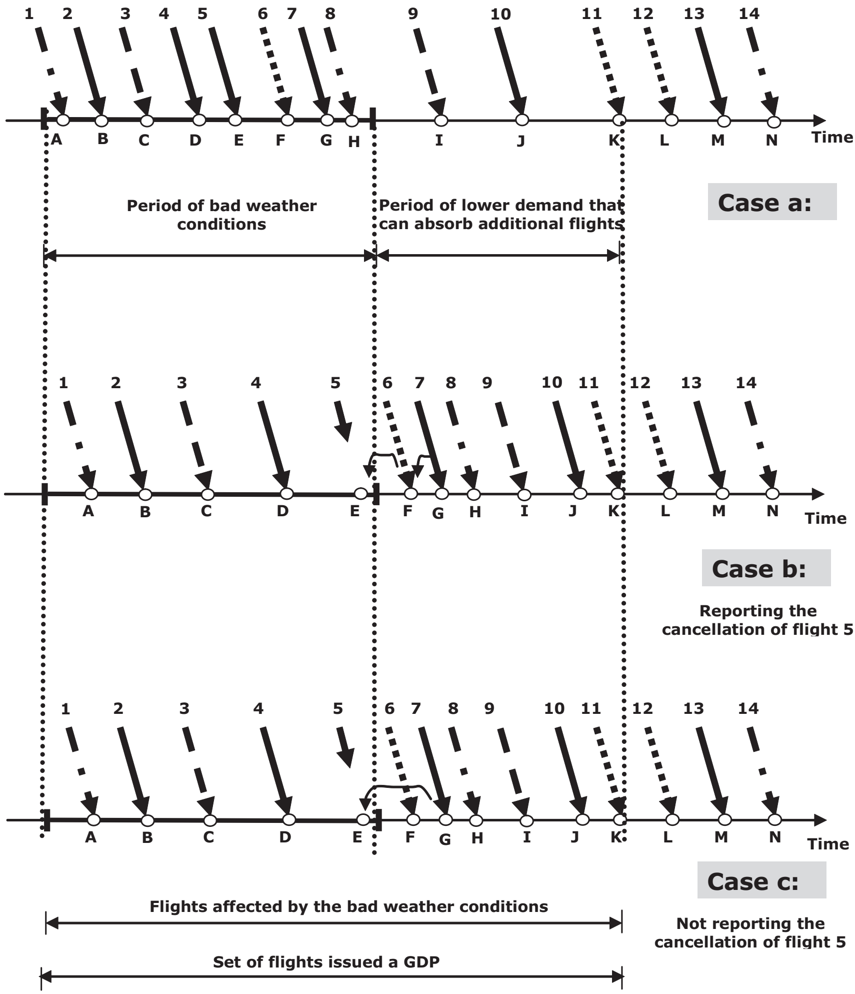


Figure 18. 4 Example of sharing information about canceling flight 5

dynamic environment. After some time has passed since the implementation the first-come-first-served rule to allocate flights to landing slots at airports during a GDP, it is concluded that there is an essential need to get real-time schedule data from the air carriers. Also, a way should be found to remove the disincentives for the air carriers to report their latest schedule information.

Collaborative Decision Making

To overcome the disincentives that preclude the air carriers from reporting their latest updated schedule in the US, a new paradigm is developed called the

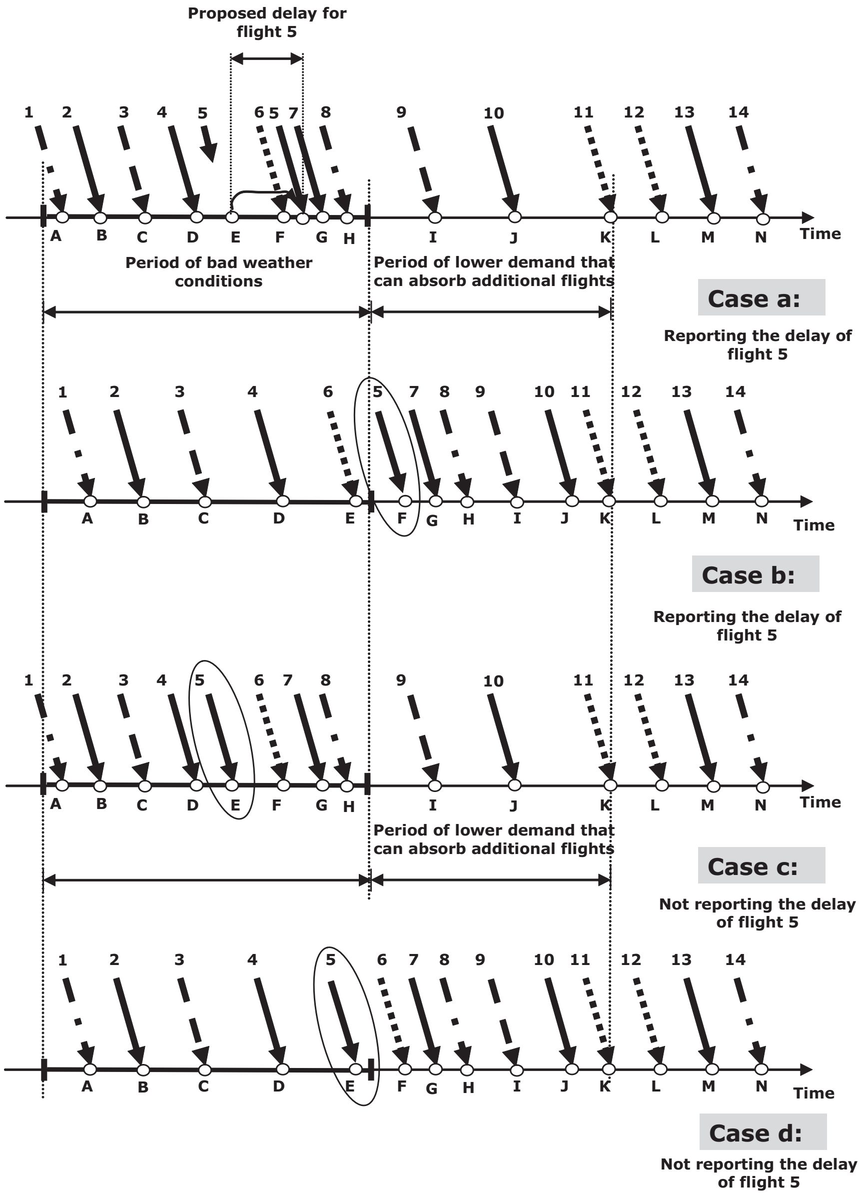


Figure 18. 5 Example of sharing information about delaying flight 5

Collaborative Decision Making (CDM) paradigm. Figure 18.6 shows the interaction between the FAA and the different air carriers under the CDM paradigm. As a first step to address the air carriers' disincentives, slots are allocated to the air carriers according to their original (static) schedule, which is known as the Ration-by-

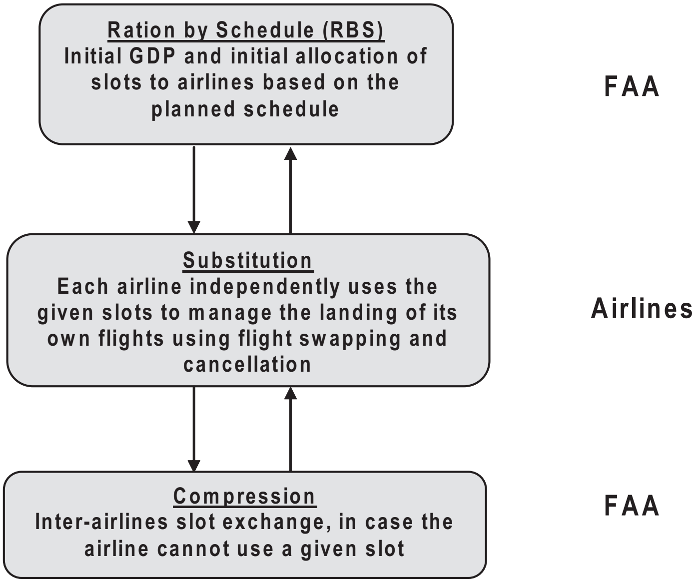


Figure 18.6 Interaction between the FAA and the airlines during a GDP under the CDM paradigm

Schedule (RBS) rule. According to the RBS, flights are prioritized by their original planned schedule, even if they are delayed or canceled. The RBS eliminates the concern of the air carrier of being penalized if they report their latest information regarding their schedule.

After the arrival slots are allocated to the air carriers, each air carrier can independently reorganize and swap its own flights into its own slots. This step, known as substitution, allows the air carrier to reduce the delay of its important flights and take advantage of the empty slots of cancelled flights. Figure 18.7 gives an example of substitution, where the air carrier that operates flights 4 and 5 exchanges the slots of these two flights. Therefore, flight 5 lands in the earlier slot D, and flight 4 lands in slot E. The air carrier can adopt this substitution if delaying flight 5 is critical to the air carrier. Another example is given in Figure 18.8, where the air carrier that operates flights 4, 5, and 7 decides to cancel flight 4 in slot D. Therefore, slot D becomes available, and the air carrier can benefit from this cancellation by moving flight 5 to slot D and also flight 7 to slot E. These two substitutions save considerable portions of the delays of flights 5 and 7 due to the GDP.

The last step in the CDM paradigm is schedule compression. To explain schedule compression, consider the GDP presented in Figure 18.9. After the landing slots are allocated to the different air carriers under the GDP, the air carrier that operates flight 4 decides to cancel flight 4. As shown in Case (b), this air carrier cannot use a substitution to move its next flight, flight 6, into slot D. To move flight 6 to slot

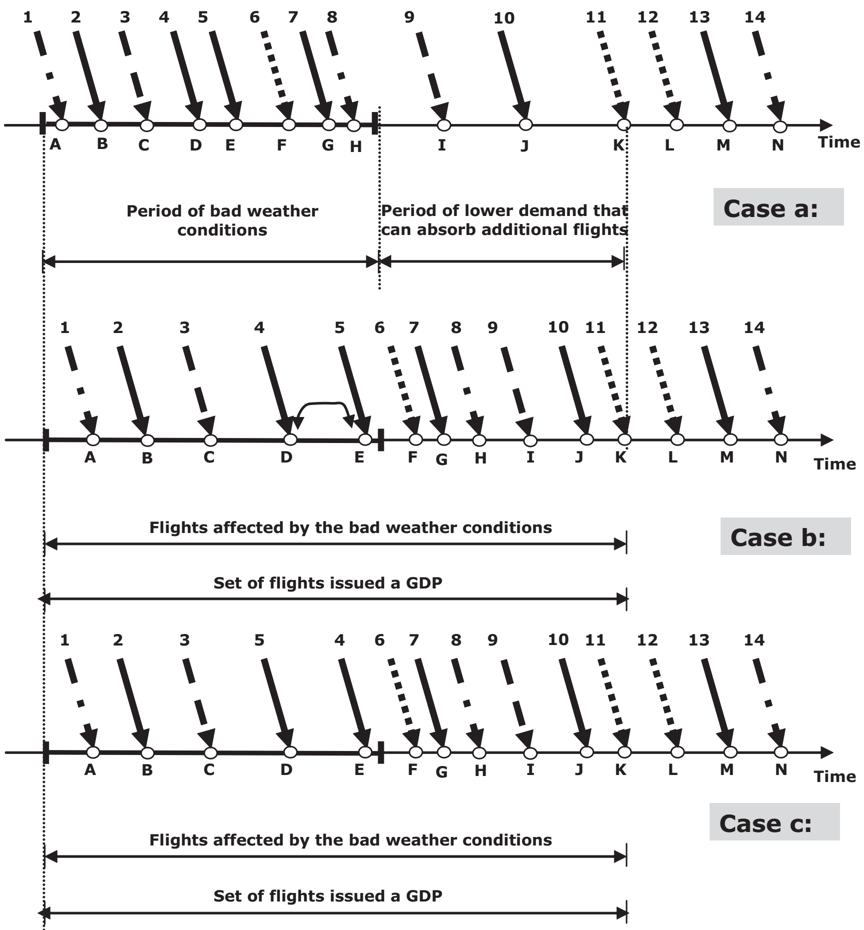


Figure 18. 7 An example of the substitution, where the air carrier that operates flights 4 and 5 exchanges the slot of these two flights

D would require that flight 6 arrive before its scheduled arrival time, which is operationally unacceptable. Regardless of its operating air carrier, the next flight is allocated to slot D. Hence, flight 5, which is operated by another air carrier, can move to slot D. This move makes slot E vacant, and flight 6 can move into it. Flight 7 can then move to slot F, and so on. Flight 10 cannot move to slot I because that would make it arrive before its scheduled arrival time, which is operationally unacceptable. Compression is a rule-based mechanism that is used to fill in landing slots that are not used by air carriers due to flight cancellation. The algorithm first tries to move a flight operated by the same air carrier into the vacated slot. If one does not exist, the slot is then opened to the next flight that can move into the slot regardless of which air carrier operates that flight. The process continues while

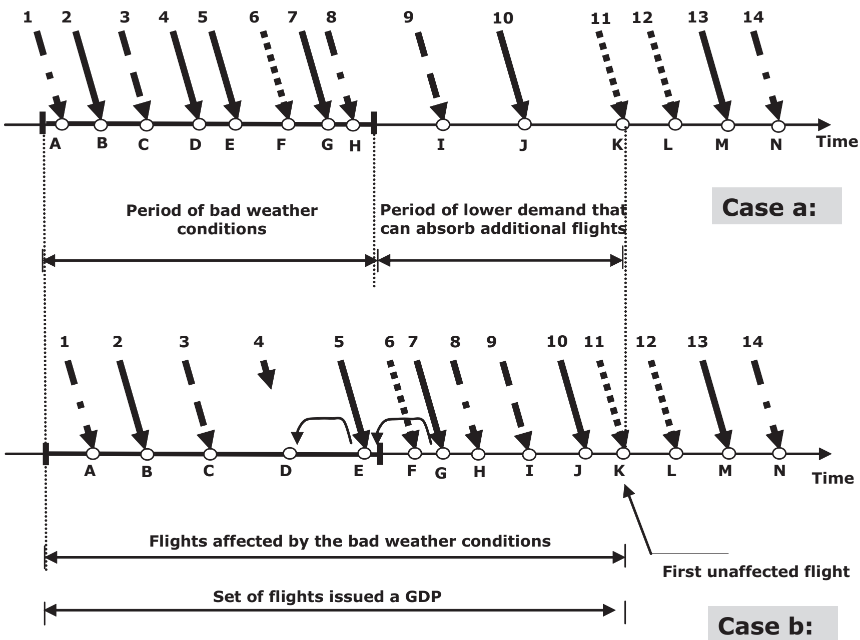


Figure 18. 8 An example of the substitution, where the air carrier that operates flights 4 and 5 cancels flight 4 to move flight 5 into the vacated slot

always checking after each slot move to see whether the air carrier that owns the first open slot can now take advantage of it.

Primary Contributions

Sokkapia (1985) addresses the issue of ground-holding strategies and proposes a very simple heuristic for the problem that essentially consists of estimating the expected delay that any flight would suffer in traveling to a congested airport then assigning a ground-holding time to this flight. The use of ground holding to resolve air traffic congestion is first described systematically by Odoni (1987). He assumes a discrete time horizon, deterministic demand, and deterministic capacity. Andreatta and Romanin-Jacur (1987) seem to be the only ones to have investigated an algorithmic approach for determining ground-holding times. They studied a simplified version of the problem with a single-time period and a limited number of flights to a single destination (airport) and developed a dynamic programming approach for obtaining an optimal ground-holding strategy to minimize the total delay costs for these flights. Terrab and Odoni (1993) present an analysis of the fundamental case in which flights from many origins must be scheduled for arrival

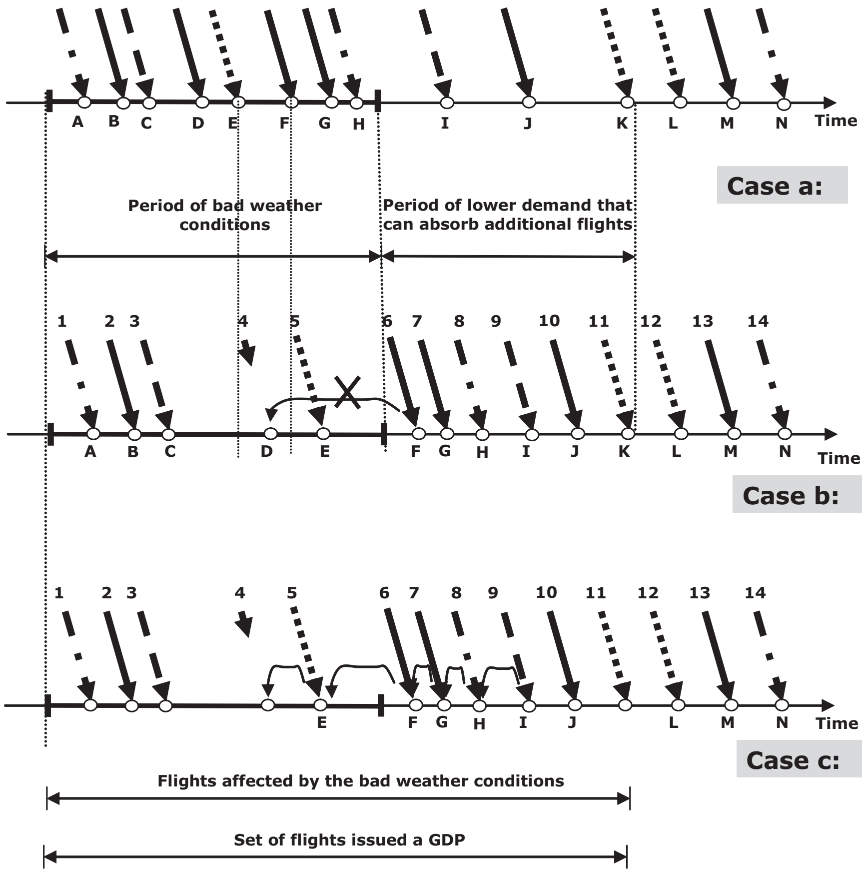


Figure 18. 9 Example of flight compression

at a single congested airport. They describe a set of approaches for addressing a deterministic and a stochastic version of the problem. Richetta and Odoni (1993) provide a linear programming solution to a multi-period single time airport case where capacity is stochastic. Richetta (1995) tests static and dynamic optimal solutions, and a very fast heuristic for the assignment of ground-holds in ATC. The optimal solutions are based on stochastic linear programming. The heuristic incorporates elements of stochastic modeling by utilizing information conveyed by a probabilistic forecast of the airport landing capacity while taking into consideration the dynamic nature of the problem. Hoffman (1997) and Hoffman and Ball (2000) present different models of the single-airport ground-holding problem with banking constraints. These constraints enforce the desire of airlines to land certain groups of flights, called banks, within fixed time windows thus preventing the propagation of delays throughout their entire operation. Bertsimas and Stock (1998) model the allocation of ground delays and also speed control of

airborne traffic. Brunetta et al. (1998) and Andreatta and Brunetta (1998) present a new heuristic algorithm that is based on priority rules of flights to solve the multi-airport ground holding problem. Carr et al. (1999) and Carr et al. (2000) propose an approach where delay exchanges in arrival sequencing and scheduling permit airlines to express relative arrival priorities so that these can be taken into account for the arrival slot allocation. This approach is proposed to be implemented instead of the traditional first-come-first-served (also known as FCFS) system.

The implementation of the ground holding with the CDM paradigm is described by Wambsganss (1996). Ball et al. (1999) provide a preliminary assessment to the CDM in air traffic management. Chang et al. (2001) present the enhancements to the FAA GDP under CDM. Ball et al. (2001) present the current and future research directions in CDM in air traffic management. An example of these directions is also presented in Ball et al. (2003), which presents an efficient solution of a stochastic ground holding problem. Also, Ball and Lulli (2004) present a methodology to optimize the GDP over the included flight set based on distance. Vossen and Ball (2006) discuss the slot trading possibility between airlines in CDM.

Considering the CDM paradigm, the problem of allocating flights to slots to reduce the downline impact of the GDP on the airline schedule (the airline's side of the problem), is first addressed by Vasquez-Marquez (1991) for American Airlines. The objective is to find a better landing-slot allocation plan after a slot becomes vacant due to a flight cancellation. The problem is formulated as Lou and Yu (1997) investigate how airlines reschedule resources to respond to the GDP. Niznik (2001) presents another model that responds to the GDP at American Airlines. He used a heuristic approach that creates multiple sets of the assignments of inbound flights that can be independently evaluated. These assignments are then combined together to create a complete solution. Rome et al (2001) are the first effort to reschedule airlines arrivals while considering the effect of downline delay. Using data from North West airlines, they attempt to minimize the downline delays by allowing the airline to swap its landing slots within ±10-minute window. However, their cost model is very simple as it considers the effect of delaying each flight independently. The model calculates each delay chain separately. Abdelghany et al. (2004) present a solution algorithm for the slot allocation problem during GDPs, considering the airlines' side of the problem. The model efficiently assigns inbound flights affected by the GDP to available landing slots such that the overall downline impact resulting from delaying any of these inbound flights is minimized. In this model, a Genetic Algorithm (GA) is used as a randomized search method. The GA is integrated with a flight simulation model. The GA searches for the optimal slot allocation pattern for the flights. The flight simulation model guides the search by evaluating the overall airline performance for each generated pattern. It captures the schedule interaction of the different resources (aircraft, pilot, and flight attendants), analyzes the downline impact of any selected slot-allocation pattern, and describes the impact quantitatively. This impact is represented in terms of the delay of flights, the misconnection of aircrafts, and the crew and passengers as well as crew illegalities.

References

Abdelghany A., Abdelghany, K.F., and Ekollu, G. 2004. A Genetic Algorithm Approach for Ground Delay Program Management: The Airlines' Side of the Problem. *Air Traffic Control Quarterly*, 12(1), 53-74.

Andreatta, G. and Brunetta, L. 1998. Multi-airport Ground Holding Problem: A Computational Evaluation of Exact Algorithms. *Operations Research*, 46(1), 57-64

Andreatta, G. and Romanin-Jacur, G. 1987. Aircraft Flow Management under Congestion. *Transportation Science*, 21, 249-253.

Ball, M., Hoffman, R., Hall, W., and Muharremoglu, A. 1999, *Collaborative Decision Making in Air Traffic Management: A Preliminary Assessment*, NEXTOR technical report RR-99-3, UC Berkeley, 1999.

Ball, M.O., Chen, C., Hoffman, R., and Vossen T. 2001. Collaborative Decision Making in Air Traffic Management: Current and Future Research Directions. In *New Concepts and Methods in Air Traffic Management*, L. Bianco, P. Dell'Olmo and A. Odoni (Eds), Springer-Verlag, Berlin, pp 17-30.

Ball, M.O., Hoffman, R., Odoni A., and Rifkin R. 2003. Efficient Solution of a Stochastic Ground Holding Problem. *Operations Research*, 51, 167-171.

Ball, M.O. and Lulli, G. 2004. Ground Delay Programs: Optimizing over the Included Flight Set Based on Distance. *Air Traffic Control* Quarterly, 12, 1-25

Bertsimas, D. and Stock, S. 1998. The Air Traffic Flow Management Problem with Enroute Capacities. *Operations Research*, 46, 406-422.

Brunetta, L., Guastalla, G., and Navazio. L. 1998. Solving the Multi-airport Ground Holding Problem. *Annals of Operations Research*, 81(1), 271-287.

Carr, G.C., Erzberger, H., and Neuman, F. 1999. Delay Exchanges in Arrival Sequencing and Scheduling. *Journal of Aircraft*, 36, 785-791.

Carr, G.C., Erzberger, H., and Neuman, F. 2000. Fast-time Study of Airline-influenced Arrival Sequencing and Scheduling. *Journal of Guidance, Control and Dynamics*, 23, 526-531.

Chang, K., Howard, K., Oiesen, R., Shisler, L., Tanino, M., and Wambsganss, M. 2001. Enhancements to the FAA Ground-delay Program under Collaborative Decision Making. *Interfaces*, 31(1), 57-76.

Hoffman, R.L. 1997. *Integer Programming Models for Ground-Holding in Air Traffic Flow Management*, PhD dissertation, University of Maryland, 1997.

Hoffman, R. and Ball, M.O. 2000. A Comparison of Formulations for the Single Airport Ground Holding Problem with Banking Constraints. *Operations Research*, 48, 578-590.

Luo, S. and Yu, G. 1997. On the Airline Schedule Perturbation Problem Caused by the Ground Delay Program. *Transportation Science*, 31, 298-311.

Niznik, T.J. 2001. Optimizing the Airline Response to Ground Delay Programs. *AGIFORS* 2001, *Airline Operations*, Ocho Rios, Jamaica.

Odoni, A. 1987. The Flow Management Problem in Air Traffic Control. In A. Odoni, L. Bianco, G. Szego, (Eds) *Flow Control of Congested Networks*. Springer-Verlag, Berlin, Germany, 269-288.

Richetta. O. 1995. Optimal Algorithms and a Remarkably Efficient Heuristic for the Ground-holding Problem in Air Traffic Control. *Operations Research*, 43(5), 758-770.

Richetta. O. and Odoni, A.R. 1993. Solving Optimally the Static Ground Holding Policy Problem in Air Traffic Control, *Transportation Science*, 24, 228-238.

Rome, J., Rose, S., Lee, R., Cistone, J., Bell, G., and Leber, W. 2001. Ripple Delay and its Mitigation. *Air Traffic Control Quarterly*, 9 (2), 59-98.

Sokkapia, B.G. 1985. *Arrival Flow Management as a Feedback Control System*. The Mitre Corporation, Washington, DC.

Terrab, M., and Odoni, R. 1993. Strategic Flow Management for Air Traffic Control. *Operations Research*, 41(1), 138-152.

Vazquez-Marquez, A. 1991. American Airlines Arrival Slot Allocation System (ASAS). *Interfaces*, 21(1), 42-61.

Vossen, T. and Ball, M.O. 2006. Slot Trading Opportunities in Collaborative Ground Delay Programs. *Transportation Science,* 40(1), 29-43.

Wambsganss, M.C. 1996. Collaborative Decision Making Through Dynamic Information Transfer. *Air Traffic Control Quarterly*, 4, 107-123.

This page has been left blank intentionally

Chapter 19
Impact of Disruptions on Air Carrier Schedule

Airline Resources

In the previous chapter, we saw how an air carrier can be forced to delay or cancel one or more of its flights due to GDPs. Other factors might also result in flight delays or cancellations including aircraft break, security breaches, late crew, and so on. In this chapter, the possible impact of flight delay and cancellation on the air carrier schedule is investigated.

Aircraft

Aircraft is the most valuable resource for the air carrier. Air carriers tend to maximize the utilization of each aircraft by maximizing the number of its daily flights. This is typically achieved by minimizing the ground time that an aircraft spends at airports between two successive flights. Any disruption, delay, or cancellation in the aircraft schedule will certainly have a downline impact on other flights. For example, consider the hypothetical aircraft route given in Figure 19.1. The aircraft route extends over four days, and starts and ends at San Francisco (SFO). The origin and destination of each flight are given by the airport code. This aircraft is planned to remain overnight at DTW, then at DEN, and finally at SEA. There is a planned maintenance activity while the aircraft is staying overnight at DEN. Assume that the flight LAX-ORD, which is scheduled on the second day of this aircraft route, has to be delayed, as shown in Figure 19.1. This delay could be due to a GDP at ORD or any other reason. If the delay is long enough, it might cause the next flight ORD-DEN to be delayed. The scheduled maintenance of the aircraft at DEN might also be affected because of this delay. Furthermore, if the flight LAX-ORD is to be canceled, the aircraft would be stranded at LAX, and there would be no aircraft at ORD to operate the next flight, ORD-DEN. Also, the aircraft would not be available for its maintenance at DEN.

Crew

As mentioned earlier, there are about 12 different groups of employees that work together to operate each flight in the air carrier schedule. These groups includes pilots, cabin crew (flight attendants), maintenance crew, ramp agents, baggage handlers, cargo agents, fueling agents, customer service agents, gate agents,

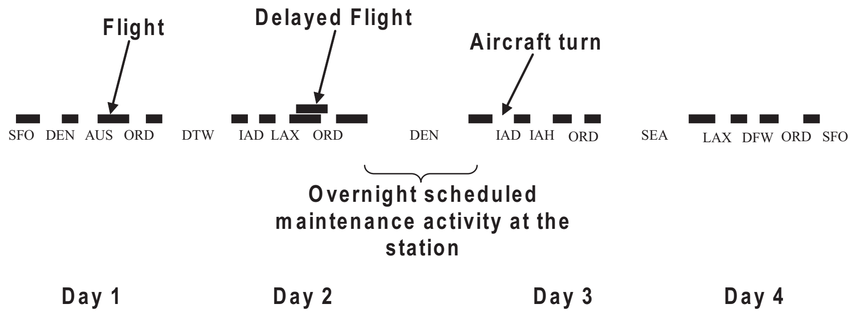


Figure 19.1 A hypothetical aircraft route

catering agents, aircraft cleaning agents, and operations agents. To maximize the total throughput of labor, most of these employees operate on a tight schedule of successive flights. Any disruption of schedule of any of these groups might affect their downline schedule and result in the risk of violating their work rules.

Pilots and flight attendants are the most critical groups because they are not positioned at one airport. Their schedule is designed such that they travel from one airport to another serving the different flights. As noted earlier, pilots and flight attendants are scheduled to travel in trippairs. Each trippair extends over 2–7 days (duty periods). Each duty period consists of a set of successive flights and are separated by crew rest periods. Figure 19.2 shows an example of a trippair that has three duty periods and two layovers. The airport code of the origin and destination of each flight are also given. It should be noted that the operation rules related to the minimum length of the connection period and the rest period, as well as the maximum length of the duty period, are set by the aviation administration to make sure that the crew are getting enough rest and not exhausted during their work. Air carriers and pilots might negotiate more relaxed operational rules in their contract to improve the quality of life for the crew.

Three different problems could happen to the crew schedule if the schedule of one or more flights is disrupted. These problems can be defined as follows:

- *Misconnect Problem*. This problem occurs when a connecting crew member arrives late and is unable to fly the next flight in the same duty period on time. Figure 19.3 shows an example of a typical misconnect problem in a trippair. When flight DEN-AUS is delayed, the new connection time becomes less than the minimum required connection time for the crew. Therefore, another crew member must be found at AUS for the next flight AUS-SFO to depart on time.
- *Rest Problem*. This problem occurs when a crew member has a rest period (layover) that is less than the minimum required rest period. The problem could result from a late arrival at the end of the previous duty period. In this case, the crew member would be unable to fly the first segment of the next duty period on time. This problem is a special case of the misconnect problem that

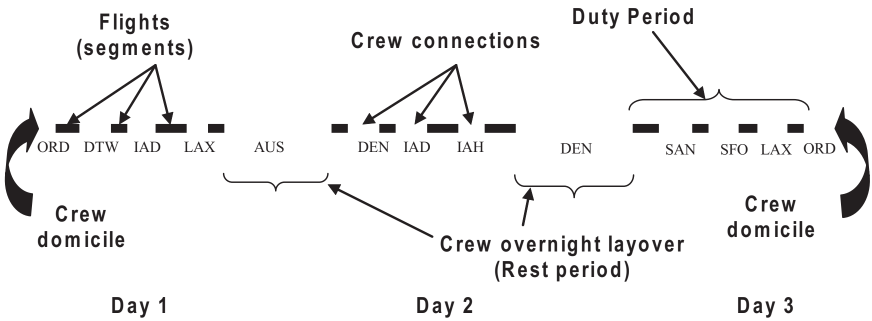


Figure 19.2 An example of a trippair that has three duty periods and two layovers

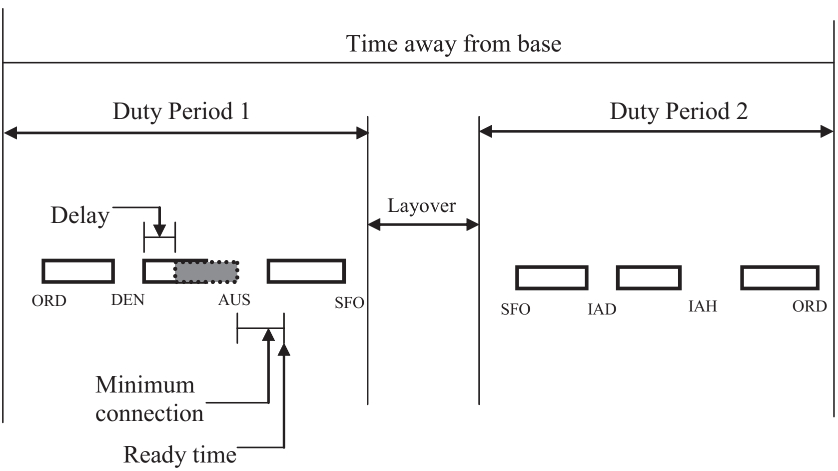


Figure 19.3 Example of a crew misconnect problem

considers the required rest period as the required connection time. Figure 19.4 shows a typical example of the rest problem. When flight AUS-SFO is delayed, the layover becomes less than the minimum required rest period. The first flight of the next duty period, SFO-IAD, cannot depart on time with the originally assigned crew because the crew must have the minimum legal rest at SFO.

- *Duty Problem*. This problem occurs when a duty period limit is exceeded due to a delay of one or more of the flights of this duty period. This delay could be due to ground holding, an unplanned aircraft maintenance, a longer than expected customer service time, and so on. In this case, the crew cannot fly the last flight of the duty period because of this duty problem. Figure 19.5 shows an example of a duty problem. Flight AUS-SFO is delayed, and its arrival time passes the duty limit of the crew. Therefore, another crew must be found at station AUS to fly flight AUS-SFO.

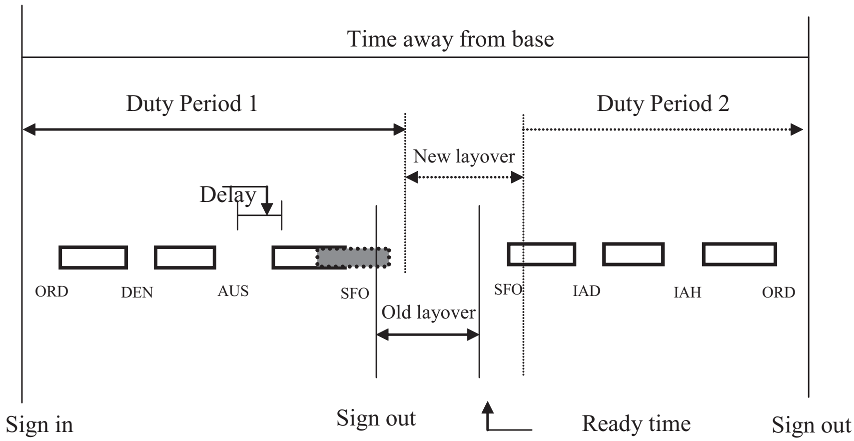


Figure 19.4 Example of crew rest (layover) problem

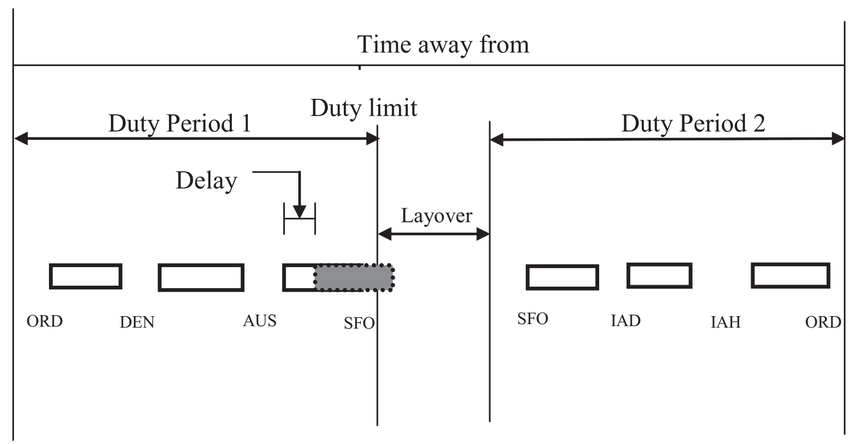


Figure 19.5 Example of crew duty problem

Slack Time

Slack Time of a Resource

In the previous section, the structure of the schedule of the main air carrier resources, aircraft, and crew, are presented. Typically, these resources are scheduled to operate a set of consecutive flights. A delay of a resource on one flight in its schedule will lead to a delay of one or more of its successive flights. After the arrival time of an aircraft or a crew member on a flight, this aircraft or the crew member is allowed some time to get ready for the next flight. For example, for an aircraft, this time is necessary for passengers deplaning, unloading cargo and luggage, cleaning the aircraft, refueling, loading cargo and luggage for the next flight, boarding

passengers of the next flight, and so on. Similarly for a crew member, this time is required to walk from the gate of the inbound flight to the gate of the next flight. If the resource is ready for the next flight before its scheduled departure time, there is a slack time for this resource. The slack time of the resource is defined as the difference between the departure time of the outbound flight and the ready time of the resource. Figure 19.6 shows an example of the slack time of a resource connecting between flights F1 and F2, which are represented by two rectangles. After the resource arrives on flight F1 at A1, a predefined period of time is given to this resource to get ready for the next flight. The slack time is the difference between the departure time of F2, D2, and the ready time of the resource.

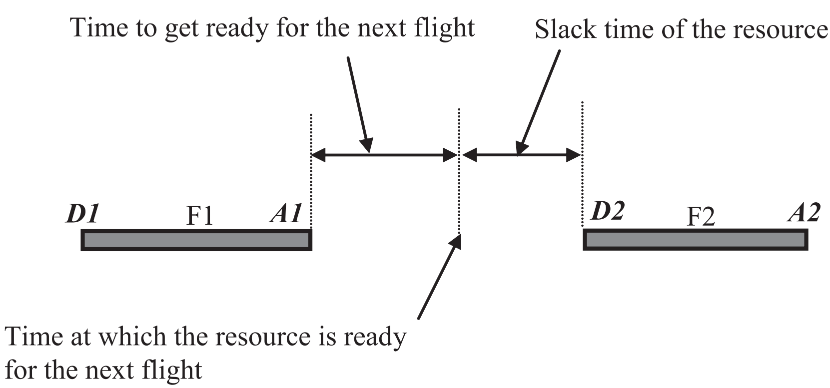


Figure 19.6 Example of a slack time of a resource

The slack time represents an idle time for the resource. The longer the slack time of a resource, the less the number of the flights the resource can operate. Therefore, air carriers are trying to cut slack times to maximize the productivity of their resources. However, the slack time has the advantage that it can absorb and alleviate the impact of schedule disruption. For example, if flight F1 is delayed as shown Figure 19.7, flight F2 will not be affected as long as the delay of flight F1 is less the slack time of the resource. Flight F2 will be delayed if the delay of flight F1 is more than the slack time of the resource. Therefore, the slack times of resources are important to increase the schedule robustness and reduce the chances of disruptions during adverse weather conditions. However, the best trade-off should be found between the slack times and resources productivity.

Slack Time of a Flight

Typically in the air carrier schedule, a flight feeds its resources to one or more downline flights. We define the slack time of a flight as the maximum amount of time this flight can be delayed without affecting any of its downline flights. Mathematically, slack time is calculated as the least amount of slack time of all the outbound resources of this flight. To explain the slack time of a flight, consider

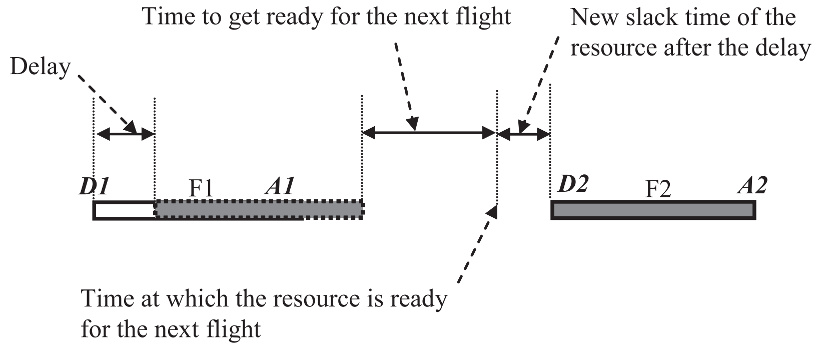


Figure 19.7 Resource slack with flight delay

flight F1 in Figure 19.8, which departs at D1 and arrives at A1. Upon arrival of flight F1 at its destination, it feeds resources such as aircraft and crew to other flights. Assume that a pilot, R_{F1F3}, is connecting from flight F1 to flight F3 and the aircraft, R_{F1F4}, is connecting from flight F1 to flight F4. The pilot, R_{F1F3}, has a slack time of S_{F1F3} and the aircraft, R_{F1F4}, has a slack time of S_{F1F4}, as shown in Figure 19.8. The slack time of flight F1 is the minimum time between S_{F1F3} and S_{F1F4}.

Network Connectivity

The schedules of the different aircraft and crew members are put together to form the total schedule of the air carrier. The schedules of the different resources are very interdependent and connected. To explain the resources connectivity, consider the hypothetical air carrier network given in Figure 19.9. Each rectangle represents

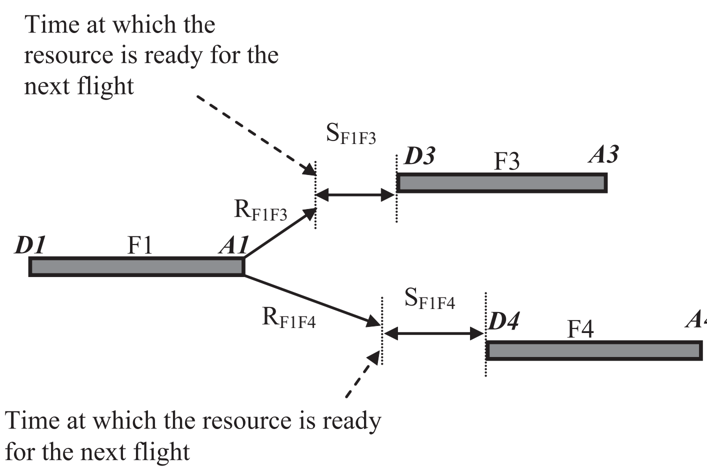


Figure 19.8 Representation of the flight slack time

a flight where the start and the end of the rectangle represents its departure and arrival times, respectively. The arrows represent the connecting resources (aircraft and onboard crew) into and out of flights. One arrow represents one resource. The start of the arrow is the arrival time of the inbound flight. The end of the arrow is the resource ready time to operate the next flight. Assume also that flights F1 and F2 arrive at the same airport and feed resources to flights F3, F4, and F5, which depart from the destination of flights F1 and F2. Flight F1 departs at D1 and arrives at A1. Upon arrival of flight F1, resource R_{F1F3} connects to flight F3, and resource R_{F1F4} connects to flight F4, respectively. Flight F2 departs at D2 and arrives at A2. Upon its arrival, resource R_{F2F4} connects to flight F4, and resource R_{F2F5} connects to flight F5, respectively. Also, for simplicity, assume that flights F1 and F2 are the only flights that feed resources to flight F4. In the normal operation conditions, all resources are planned to be ready before the scheduled departure time of their next assigned flights by some time called the slack time *S*. A resource slack time is the difference between the resource next flight departure time and its ready time. As shown in Figure 19.9, resources R_{F1F3} and R_{F1F4} have the positive slack times of S_{F1F3} and S_{F1F4}, respectively. If a resource is delayed within its slack, its downline flight would still depart on time. If S_{F1F3} is less than S_{F1F4}, flight F1 could be delayed S_{F1F3} minutes without affecting any downline flights. This defines a flight *downline slack*, which is the longest time interval a flight can be delayed without affecting any of its downline flights. Computationally, a flight downline slack is the shortest slack among all its outbound resources.

Assume a hypothetical decision of delaying flight F1 such that it departs at D1′. This delay could be a result of a GDP imposed by the FAA at the arrival station of flight F1, or a result of a plan to reassign flights to landing slots by the air carrier management, or any other reason. For simplicity, assume that flight F2 is not delayed,

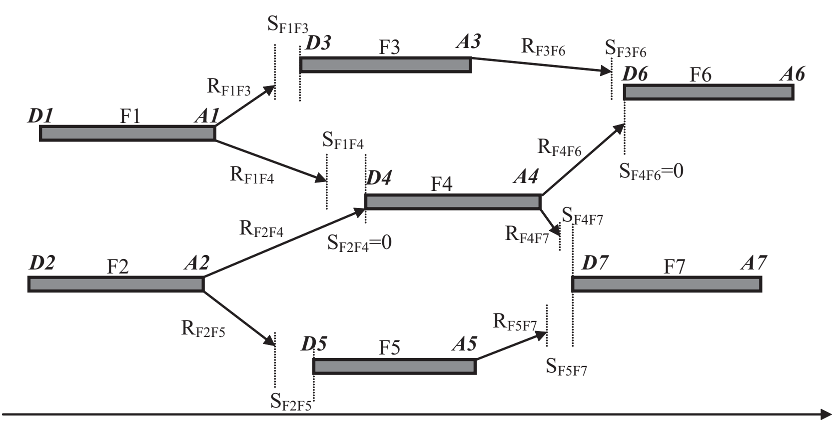


Figure 19.9 Connectivity under normal operation conditions

as shown in Figure 19.10. If the delay of flight F1, (D1′–D1), is greater than the downline slack of flight F1, one or more of the downline flights that use resources out of flight F1 would be affected. As shown in Figure 19.10, the delay (D1′–D1) is greater than the slack time of the connecting resources, causing these resources (and passengers) to misconnect flights F3 and F4. Thus, if no recovery action is considered, flights F3 and F4 would be delayed until the arrival of their resources from flight F1. If substitutes are found for the resources and flight F4 to depart on time, connecting passengers from flight F1 to F4 would miss their connection. Therefore, these passengers have to be reaccommodated on different flights.

For flight F4, assume the inbound resource R_{F1F4} is an aircraft while the inbound resource R_{F2F4} is a flight attendant. Figure 19.10 shows the end of the maximum duty period of this flight attendant. When flight F1 is delayed and there is no substitute available for the aircraft R_{F1F4}, flight F4 must wait for its aircraft $R_{F1F4,}$ and will also be delayed. This delay will cause a duty limit violation for flight attendant R_{F2F4}. A duty limit violation occurs for a flight attendant when they work more than the predefined maximum duty period. Therefore, this flight attendant cannot serve on flight F4. Flight F4 has to be canceled or another flight attendant must be made available to operate flight F4 instead of R_{F2F4}. Additionally, delaying flight F4 causes its outbound resources to misconnect their next flights. Resources R_{F4F6} and R_{F4F7} would misconnect flights F6 and F7, respectively. Also, delaying flight F4 would cause a workload disruption to all its supporting airport facilities and staff including gates, baggage handling facilities, customer service agents, ground staff, and so on. Therefore, it is obvious that delaying flight F1 impacts the decisions of delay, cancellation, and resources recovery for flights F3, F4, F6, and F7.

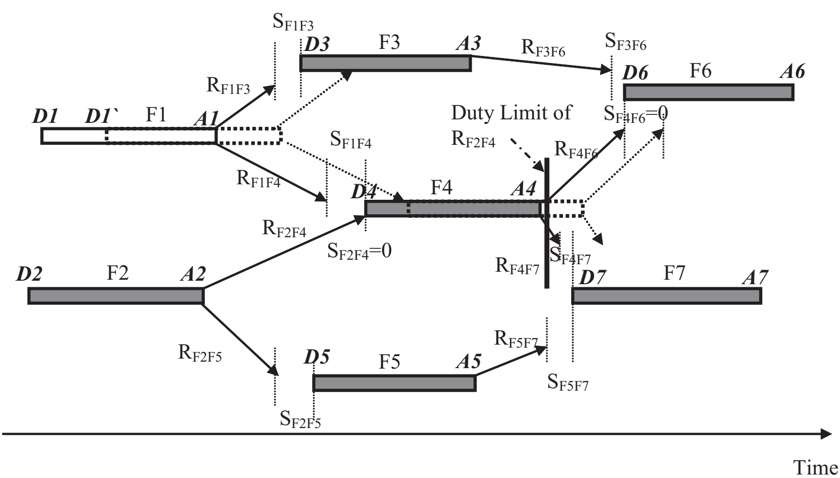


Figure 19.10 Resource connectivity after delaying flight F1

This example demonstrates the intensive connectivity among the flights' operating resources and the importance of considering the downline impact when evaluating a change in flight schedule. The example illustrates the snowball effect that could result from changing the schedule of just one flight in a small and simple air carrier network. Identifying the set of downline impacted flights due to changing the schedule of a flight and evaluating this impact could require a considerable computation effort. Also, it requires an awareness of all the operational rules and legalities of the different flight resources so that none of these rules is violated. In addition, the example shows the relation between the main operating resources of flights only. However, the same logic can be extended to show the impact of flight delay on connecting passengers. Furthermore, the example also presents the interaction between the main decisions in the recovery process. These decisions are related to modifying flights schedule (delay, cancellation, and diversion) and workloads of affected resources (aircrafts, airport facilities, operating crews, and supporting staff) as well as reaccommodating any stranded passengers. For example, the decision of delaying flight F1 affects the decision of delaying, canceling, or recovering the resources of flights F4, F6, and F7. Going backward, if it is known that flights F4, F6, or F7 cannot be delayed or cancelled, the decision to delay flight F1 should be avoided.

Primary Contributions

Investigating the impact of flight delay on the airline schedule is considered by Beatty et al. (1999) who investigate the impact of individual flight delays as they propagate by simulating the movement of delayed cockpit crews, flight attendants, and aircraft through the network. The relationship between the time of day and delay propagation is also investigated. They use a metric called delay multiplier to measure the ratio of propagated delay minutes with respect to the initial delay duration. Wang et al. (2003) use queuing models to present a simple analytical model that explicitly separates the controllable factors that influence delays and the propagation of delays in the National Airspace System (NAS) from those factors that are random variables in a given scenario. They show how the model can be applied from specific NAS airports to better understand delay propagation, especially the effects of flight schedule parameters on measured delay.

Abdelghany et al. (2004) present a flight delay projection model that projects flight delays and alerts for downline operation breaks for large-scale airline schedules. In this model, the airlines' daily schedule is represented in the form of a directed acyclic graph with its nodes being the different scheduled events, and its arcs being the scheduled activities between these events. Scheduled events include flight departures, flight arrivals, crew duty starts, crew releases from duty, aircraft maintenance starts and ends, and so on. The in-between activities include aircraft taxi-out and taxi-in, flying, crew connections, layovers, and so on. Using this graph representation, the model applies the classical shortest path algorithm

to determine the earliest possible time at which the different events could occur while considering all operation constraints that govern the operation including the GDP issued by the FAA and crew legality rules. The results of applying the model at the operation control center of a major airlines company in the US are presented. Janic (2005) presents a model for the assessment of the economic consequences of the large-scale disruptions of an airline with a hub-and-spoke network. These consequences are expressed by the cost of delayed and cancelled complexes of flights. The model is based on the theory of queuing systems with the airline hub airport as a server and the complexes of flights as customers. The model is applied to the disrupted hub-and-spoke network of a large European airline.

References

Abdelghany, K., Shah S., Raina S., and Abdelghany, A. 2004. A Flight Delay Projection Model for Airline Schedule during Irregular Operation Conditions. *Journal of Air Transport Management*, 10 (6), 385-394.

Beatty, R., Hsu, R., Berry, L., and Rome, J. 1999. Preliminary Evaluation of Flight Delay Propagation through an Airline Schedule. *Air Traffic Control Quarterly,* 7 (4), 259-270.

Janic, M. 2005. Modeling the Large Scale Disruption of an Airline Network. *Journal of Transportation Engineering*, 131(4), 249-260.

Wang, P.T.R., Schaefer, L.A., and Wojcik, L.A. 2003. Flight Connections and Their Impacts on Delay Propagation. *Digital Avionics Systems Conference*, DASC '03. The 22nd, 1, p. 5.

Chapter 20
Airline Schedule Recovery

Proactive Approach

In the previous chapter, the impact of flight delay and cancellation on the air carrier schedule is investigated. We have seen that a delay or cancellation of one or more flights in the air carrier schedule can have a significant diverse impact on other downline flights in the schedule. It is important for the air carrier to determine this impact ahead of time so that possible actions can be taken to alleviate its effect on the schedule. Consider the example of the hypothetical air carrier network discussed in the previous chapter, and presented here in Figure 20.1. Assume flight F1 is scheduled to depart at 8:00 AM and is delayed for departure until 9:00 AM. Assume also that at some time before 8:00 AM, this air carrier finds out about this delay, say at 7:30 AM. At 7:30 AM, the air carrier can perform an exercise to predict the impact of delaying the departure time of flight F1 from 8:00 AM to 9:00 AM. Assume that flight F1 is feeding a pilot resource, R_{F1F3}, to flight F3, which has a slack time of S_{F1F3}. Since the pilot, R_{F1F3}, is delayed, flight F3 also has to be delayed. The amount of delay of flight F3 depends on the slack time S_{F1F3}. Assume that flight F3 is delayed for 45 minutes until the pilot, R_{F1F3}, gets ready for the flight. The new arrival time for flight F3 is 2:30 PM, as shown in Figure 20.1. Upon the arrival of flight F3, it feeds the pilot resource, R_{F3F6}, to flight F6. Since flight F3 is delayed, flight F6 is delayed while waiting for its resource, R_{F3F6}. The delay of flight F6 is shown in Figure 20.1.

It should be clear that since this exercise of tracking the air carrier resources is performed after the air carrier finds out about the delay of flight F1, say at 7:30 AM, the air carrier knows about the possible delays of flights F3 and F6 ahead of time before the departure times of these two flights. For example at 7:30 AM, the air carrier knows that flight F3, which is scheduled at 11:45 AM, is to be delayed for 45 minutes. Also, flight F6, which is scheduled at 3:30 PM, is to be delayed for 30 minutes. If this air carrier is to avoid delaying these two flights, possible actions can be investigated as early as 7:30 AM to find substitute pilots for these flights, which is known as the proactive approach. This proactive approach gives the air carriers more flexibility to make better economic and efficient decisions.

Schedule Recovery under Irregular Operations

The main objective of the air carrier schedule recovery during irregular operations conditions is to find an efficient cost-effective plan to alleviate the impact of these

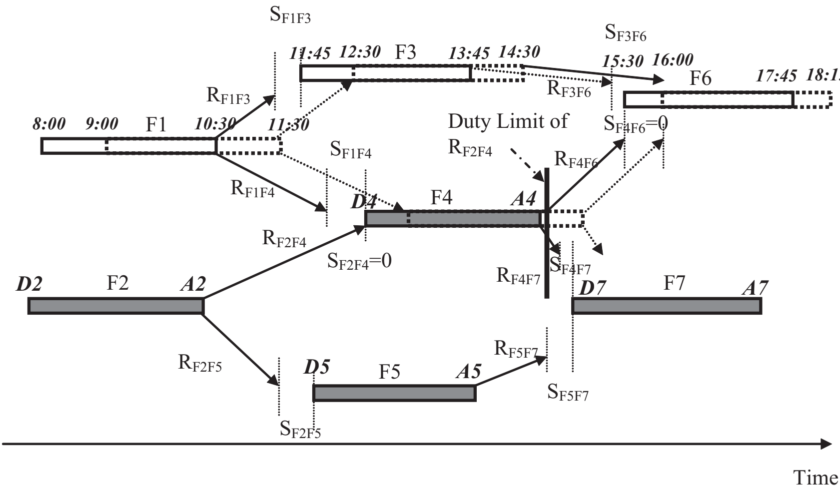


Figure 20.1 Resource connectivity after delaying flight F1

irregularities on the schedule. Five main conditions should be satisfied in any recovery plan. First, the plan should comply with any GDP issued at any airport or any air-space flow program issued to meter traffic in the space affected by adverse weather conditions. Second, this plan should also conform to regulations required for the aircraft service and maintenance activities. Third, the air carrier should meet the terms that regulate the work rules of crew and other workers. Fourth, this plan should minimize the deviation from the planned schedule by minimizing the flight delays and cancellations. Finally, the plan should consider reaccommodation of stranded passengers, crew, and aircraft. When the air carrier schedule is subjected to any source of disruption due to GDP, or any other reason and based on the size of disruption (that is, the number of affected flights), three main actions can be taken by the air carrier to recover the schedule. These actions include resource (aircraft and crew) recovery, flight delay, and flight cancellation.

Resource Recovery

The main sources of the aircraft include aircraft and crew (pilots and flight attendants).

Aircraft Recovery

Aircraft swapping: As mentioned earlier, the aircraft is the most expensive resource for the air carrier. When the schedule of the air carrier is disrupted and an aircraft is needed to operate a flight on time, it is rare to find a spare standby aircraft to operate this flight. Therefore, the only way to recover aircraft is to find a swapping opportunity with another aircraft to alleviate the impact of schedule disruptions. To explain aircraft swapping, consider the two-way swapping of

the example given in Figure 20.2, which shows parts of routes of two aircraft. The route of the first aircraft includes flights F1 and F2. The route of the second aircraft includes flights F3 and F4. As shown in the figure, flight F1 is delayed and this results in a delay for flight F2 due to the delay of the first aircraft. To avoid delaying flight F2, the second aircraft, which is available at the origin of flight F2, can be used to operate flight F2. Then, flight F4, which is scheduled later in time, can be operated by the aircraft of flight F1. This simple swap avoids the delay of flight F2. Aircraft swap typically occurs at the hub of air carrier where there are a significant number of aircraft turns that occur at the same time, and swapping opportunities can be found among them.

The two-way swap example given above can be extended to a three-way or multi-way swap as shown in Figure 20.3. Figure 20.3 shows parts of the route of three different aircraft. Aircraft I is scheduled to operate flights F1 and F2; aircraft II is scheduled to operate flights F3 and F4; and aircraft III is scheduled to operate flights F5 and F6. Similar to the previous example, Flight F1 is delayed and this results in a delay for flight F2 as a result of the delay of the aircraft I. To avoid delaying flight F2, aircraft II, which is available at the origin of flight F2 and before its scheduled departure time, can operate flight F2. However, a two-way swap cannot be made by making aircraft I operate flight F4, because it is not ready at the departure time of flight F4. Aircraft III, which is ready before the departure time of flight F4, can be used to operate flight F4. Then, aircraft I can operate flight F6, which is scheduled after the ready time of aircraft I.

Several issues should be verified when considering aircraft swapping. First, the swapped aircraft should be the same or close to the same size to accommodate passenger demand. Second, it is clear that swapping takes the aircraft out of its planned routes. In case there is any planned maintenance activity in the original route of the aircraft, the aircraft swapping prevents the aircraft from this maintenance activity. Therefore, another maintenance activity should be created

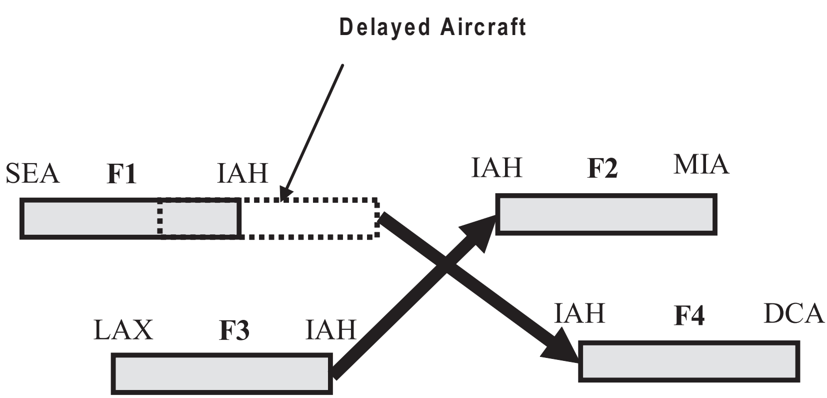


Figure 20.2 An example of two-way swap for two aircraft

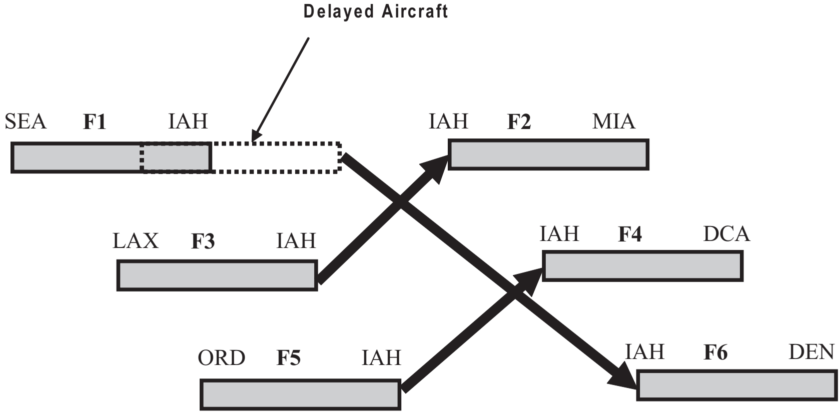


Figure 20.3 An example of three-way swap for three aircraft

for the aircraft after swapping, otherwise the proposed swap must be cancelled. When swapping aircraft, a completely new schedule should be created for each aircraft. Also, it is preferable that the aircraft returns to its original schedule, if another swapping opportunity is found as shown in Figure 20.4.

Delay to wait for aircraft: If no swapping opportunity is available between aircraft, the only other option is to delay the downline flight to wait for its inbound delayed aircraft. This delay is typically common at spokes where there are a few aircraft turns that occur at the same time. Also, the delay can be incorporated with swapping to minimize the total delay of the outbound flights.

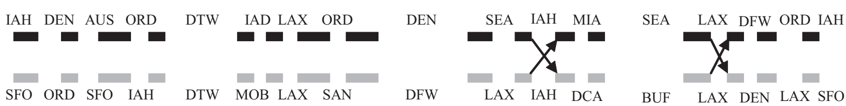


Figure 20.4 Aircraft is to return to its original schedule, if another swapping opportunity is found along its route

Crew Recovery

Standby: Air carriers typically position a predefined number of crew with different qualifications at the major stations to operate as standby crew. These crew members are scheduled to stay at the airport for a few hours (usually four hours). In case there is system disruption, these crew members may be asked to operate any of the disrupted flights. In case no flying assignment is given to them, these crew members return home by the end of their four-hour duty period. It is an interesting exercise for the air carrier to determine the number of needed standby crew members with the necessary qualifications at the needed times of the day.

Reserve crew: Reserve crew is similar to standby crew except that they are not positioned at the airport. Reserve crew members stay at home and during a

predefined period of time, they can receive phone calls for a flying assignment. The call should be at least four to five hours in advance of the given assignment, during which the crew member can get ready for the flight and arrange transportation to the airport.

Crew swapping: Crew swapping is similar to the aircraft swapping explained above. When swapping crew members, several issues should be considered. First, the swapped crew members should have the same qualifications. For pilots, the swapped pilots should be qualified to fly the same fleet and have the same position (captain or first officer). For flight attendants, the swapped flight attendants should be qualified to serve on the same fleet or aircraft type (wide body or narrow body). Second, when a swap is considered, the assignment should not violate the regulations of the crew working hours. Crew members should receive the required connection time between flights and the minimum required layover periods between duty periods, and the duration of any of their duty periods should not exceed the maximum duty period. When a crew member is taken out of their original trippair for swapping purposes, the swapping program should consider creating a new trippair for the crew member that ends at their domicile station within a few hours from the end time of their original trippair.

Stranded crew: Stranded crew members are the ones that have their outbound flight canceled and are stranded at the airport with no work assignment. These crew members can be used to fly other disrupted flights at the airports. The same issues considered when swapping crew members should also be verified when using stranded crew. If no assignment is given to stranded crew, this crew is flown to its domicile.

Deadhead: A crew member can be flown as a passenger to operate a disrupted flight at a different station. A deadhead crew member can fly on a flight for the same air carrier or a flight that belongs to another air carrier. Deadhead is also used to reposition stranded crew to its domicile.

Delay to wait for crew: Finally, if no crewmember is available to operate the disrupted flight on time, the flight can be delayed to wait for its inbound delayed crew. This option is usually considered at spoke stations where there is no substitute crew available. It is also considered when there is not enough time available to deadhead crew from other stations. Furthermore, delay is considered when the impact of this delay on other downline flights is minor.

If a resource shortage is encountered, these actions are investigated to determine their feasibility for recovering. The main challenge is to prioritize the use of these actions to solve several problems at the same time in the best efficient way. It is usually difficult to determine the impact of using a certain recovery action on the system operating cost. However, a general cost-efficient order of these actions can be as follows:

- First, a delay within an acceptable threshold (0–14 minutes) is expected to have the minimum cost impact. This delay recovers most misconnect and crew rest problems and usually has little downline effect on the system.

This small amount of delay is expected to be absorbed in the downline slacks (the difference between the actual connection time and the minimum connection time) between flights.

- Second, using the stranded resources is another inexpensive option. For example, in case of onboard crew, the crew member might be stranded due to cancellation or misconnection of their downline flights. The pay of these crewmembers is usually guaranteed, regardless of not flying the cancelled or misconnected flights. Adding different flights instead of the cancelled or misconnected ones adds no cost to the operation.
- Third, if no stranded resources are available, one can use other undisrupted (good) resources in the system for swapping purposes. Similar to the case presented in the previous example, no cost is added to the operation when changing the schedules of undisrupted resources. The only disadvantage is that those undisrupted resources (crew members) might dislike changing their original schedules, resulting in quality of life issues. Standby or reserve resources are used for the most hard-to-recover problems. The use of standby or reserve resources has to be minimized whenever possible for two main reasons: 1) using these resources for flying usually adds to the operating cost, because they have to be compensated for flying according to their contracts; and 2) schedulers should keep standby or reserve for possible unseen future resource shortage problems. Finally, longer delays or flight cancellations would be the last option available to deal with resource problems.

Sequential Recovery

The complexity of the resource recovery process is such that it ensures the recovery of three different resources; aircraft, pilot, and flight attendant, which have different availability levels and different work rules. A disrupted flight could be delayed because of the delay of more than one of its inbound resources such as the aircraft and a pilot. It is common that a substitute pilot be available to operate the flight on-time, even though there is no substitute aircraft. In this case, the flight has to be delayed anyway to wait for its aircraft.

Typically, the aircraft is the most scarce resource for the air carrier, then pilots, then the flight attendants. Generally, air carriers follow a sequential approach to recover any disrupted flight. The recovery process tries first to secure an aircraft for the flight. If an aircraft is found, the process continues to find the pilots, then the flight attendants. If one of the resources is unavailable, the flight is delayed or canceled, and the spare resources are used to operate other flights. Figure 20.5 shows a hierarchy for the sequential recovery approach. It should be mentioned that the recovery process can be performed on one flight at a time or simultaneously to a group of flights. This latter approach requires a trade-off between the different actions that can be taken to the different flights, where, for example, a spare resource can be allocated to one flight and not to the other.

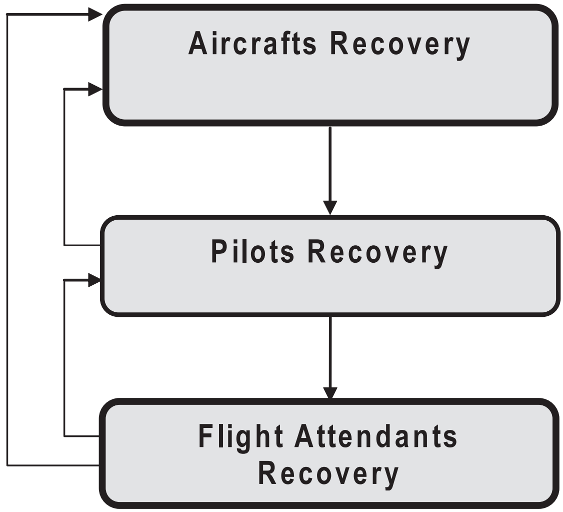


Figure 20.5 Hierarchy for the sequential recovery approach

The size of the recovery problem depends on the number of flights affected by the bad weather conditions. The recovery process is either performed manually or automatically using computer software. Most major air carriers with large numbers of daily flights automate the recovery process using computer software that differs in its methodology and capability.

Most air carriers follow the sequential approach where the recovery process is performed through three different desks in the air carrier operation management center. These desks are the aircraft recovery desk, the pilot recovery desk, and the flight attendants recovery desk. The decisions of these desks are connected to each other by a fourth group, which is called the controllers, as shown in Figure 20.6.

Flight Delay as a Recovery Action

As discussed in Chapter 17, the collaborative decision making (CDM) process between the FAA and the US air carriers gives each air carrier the flexibility to reassign its flights that are affected by the GDP to its own slots to reduce the downline impact of the GDP on its schedule, if possible. The air carrier has the flexibility to reassign its flights in its own slot to reduce the impact of the GDP on the important flights. For example, in Figure 20.7, the air carrier that operates flights 4 and 5 and is given slots D and E can swap the landing of these two flights to reduce the delay of flight 5 and alleviate the adverse consequences resulting from its delay. The objective of the air carrier is to find the optimal allocation of flights to the slots given by the FAA in order to reduce the impact of the GDP on its

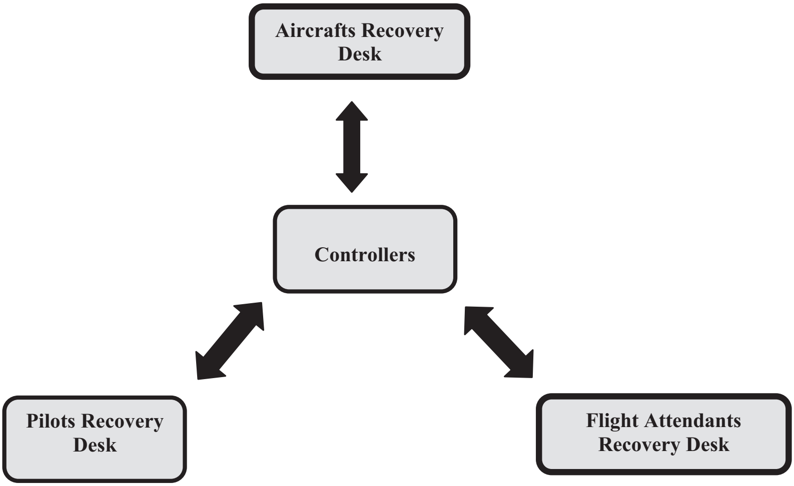


Figure 20.6 The relationships between the main recovery desks in the airline's operations center

downline schedule. Developing the optimal landing-slot allocation plan for flights during GDP is not an easy task because of three main reasons.

1. the difficulty involved in evaluating the impact (gains and losses) of moving a flight to a different slot (a schedule change of a flight);
2. the complexity of the business rules and constraints that the air carrier should follow while rescheduling its flights; and
3. the large size of the problem in terms of the number of feasible slot allocation patterns that could be generated.

The following sections clarify these reasons in more detail.

Evaluating the Impact of Schedule Change

To illustrate the difficulty involved in evaluating the gains and losses that result from changing the schedule of flights, consider the hypothetical flight connections given in Figure 20.1. In normal operating conditions, all resources are planned to be ready before the scheduled departure time of their next assigned flights by some time called the slack time *S*. When flight F1 is delayed for (D1′–D1), and this delay is greater than the downline slack of flight F1, one or more of the downline flights that use resources out of flight F1 are affected. As explained in the previous chapter, the delay of flight F1 causes the delay of flights F3 and F4. It also causes a disruption to the resources connecting from flight F2 and possible disruptions to downline flights F6 and F7. This example demonstrates the intensive connectivity among the air carrier resources, and the importance of considering the whole

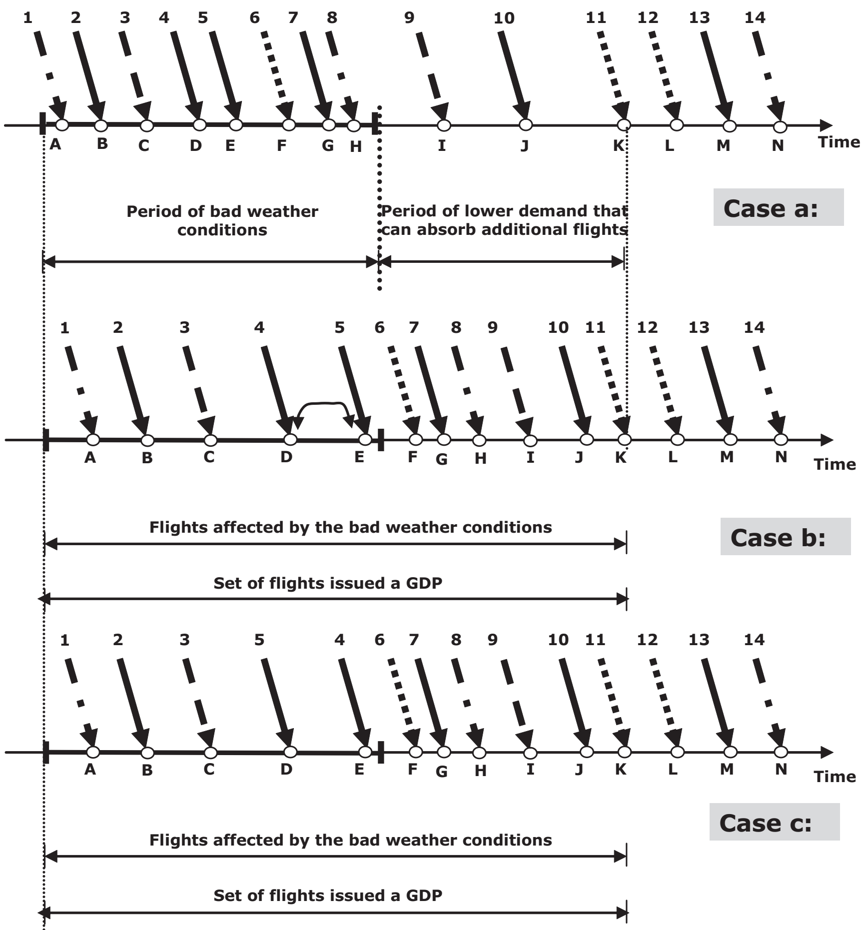


Figure 20.7 Flight delay as a recovery action

network while evaluating a change in a flight schedule rather than considering the first-level downline impact. It illustrates the snowball effect that could result from changing the schedule of just one flight in a small simple air carrier network. Determining the set of downline impacted flights due to changing the schedule of a flight and evaluating this impact could require a considerable computation. Also, it requires an awareness of all operational rules and legalities of the different flight resources such that none of these rules is violated.

To complete our discussion of this example, it should be noted that the system downline impact could be different if flight F1 is assigned to any other slot with a different delay. Also, in this example, it is assumed that flight F2 is not delayed. If flight F2 is allocated to any other slot, the downline impact of the delay of flight F1 could change. Furthermore, it is assumed that the schedule of all downline

flights of flight F1 is fixed. Any of these flights could be subjected to a change in its schedule as well, which would result in a different downline impact on the system.

Complexity of the Business Rules and Operational Constraints

Air carriers usually maintain a set of business rules and operating constraints in the event of implementing a schedule change due to GDP. One important constraint is that a flight cannot depart before its scheduled departure time. All arrival slots programmed before the scheduled arrival time of a flight cannot be considered for this flight. For example, in Figure 20.7, flight 5 cannot be assigned to slot B because it is before its scheduled arrival time. This sets a lower limit on the slots that a flight can be allocated to while solving the problem. Also, air carriers set an upper bound on the amount of delay assigned for a flight during GDP, which also limits the slots that a flight can be allocated to. For example, in Figure 20.7, by considering the maximum allowed delay for flight 5, flight 5 can be allocated to the slots J or M. According to the upper and lower limits of slots that can be allocated to flight 5, this flight can be allocated to only three possible slots (D, E, and G). In addition, air carriers schedule their flights to arrive in time banks (especially for air carriers using the hub-spoke network structure). Flights from the same origin are usually designed to be in different time banks. This scheduling introduces a minimum gap between flights arriving from the same origin. Accordingly, in solving a GDP problem, a sequence of flights that arrive from the same origin airport should remain in the same order in the new reallocation. This ordering implies that the reallocation of a flight could affect all of the subsequent flights that arrive from the same origin airport. For example, in Figure 20.7, given that flights 4, 5, 7, and 10 are arriving from the same origin, these flights have to be assigned to slots that maintain their order in the new pattern. For instance, if a solution is obtained such that flight 5 remains in slot E, flight 7 can only be assigned to slot F or any slot later than slot F that satisfies its maximum delay condition. If flight 5 is moved to the slot of flight D, flight 7 would have more slots for reallocation (slot E or any slot later than slot E that satisfies the maximum delay condition).

Too Many Patterns to Evaluate

It is well known that the problem of the landing-slot allocation is NP-hard, where the number of flight-slot allocation patterns increase exponentially as the number affected by the GDP increases. For a large-scale commercial air carrier with one of its major hubs subject to a GDP, one could reasonably assume that close to 100 flights could be affected by the GDP. If there are 50 feasible slots for each flight, the number of flight-slot allocation patterns that need to be evaluated could reach 9e+128. Ideally, obtaining the optimal solution would require evaluating all possible landing-slot allocation patterns and selecting the pattern with the best system performance, which is practically not feasible.

Flight Cancellation as a Recovery Action

Flight cancellation is usually an unfavorable decision for schedule recovery during irregular operation conditions, as it represents the most costly decision during the process. However, flight cancellation decisions are sometimes unavoidable during severe adverse weather conditions that last for long periods of the day. Canceling a flight due to adverse weather condition depends on four main factors: 1) Unavailability of landing slots within the acceptable flight delay range. During severe adverse weather conditions, air carriers usually get fewer slots than the number of their scheduled flights or in the best cases get landing slots that are very late in the day after the weather conditions improve. In this case, flight cancellation is unavoidable. 2) Lack of one of its operating resources. A resource such as aircraft or crew might not be available to the timely operation of a flight due to a disruption in its upstream schedule. If this resource is not replaceable, its corresponding flight might need to be cancelled. 3) Flight profitability represented by the number of its local and connecting passengers. Air carrier management usually tries to avoid canceling profitable flights that have a large number of local and connecting passengers. This consideration alleviates the adverse impact on as many passengers as possible and reduces the effort that is needed to re-book stranded passengers on its schedule. 4) The downline impact of cancellation on the air carrier schedule. The downline impact of flights that are candidates for cancellation is considered. Cancellation is usually avoided for flights whose operating resources have a significant downline workload in order to reduce the downline impact of the adverse weather conditions.

When adverse weather conditions occur at one of the major airports and extend over a long period of time, most air carriers gets fewer slots than they plan to get during normal operations conditions. Also, most flights get longer periods of ground delay. In that case, the air carriers are forced to cancel some of their flights. For example, in Figure 20.8, only ten slots are created during this GDP and accordingly, flights 11, 12, 13, and 14 have to be canceled.

Air carriers may also cancel a flight in an earlier slot to open this slot for another important flight that would otherwise get a significant ground delay. For example, in Figure 20.9, the air carrier that operates flights 4, 5, and 7 cancels flight 4. By canceling, slot D becomes open to flight 5, and when flight 5 is moved to slot D, flight 7 can use slot E, which saves the delay of flights 5 and 7, respectively.

Selecting which flight to cancel is typically not straightforward as the air carrier has to evaluate the revenue loss due to the cancellation, the cost of reaccommodating passengers, the ease of repositioning the resources of the cancelled flights, the scheduling of maintenance, the availability of landing slots, and so on. Special attention is also given to aircraft that might be stranded at locations with very adverse weather conditions and damaging impact, such as hurricanes, hails, snow storms, and so on. To explain this situation, consider the aircraft route shown in Figure 20.13, where the ORD-SEA has to be cancelled. To reposition the aircraft, two options are available to the air carrier. First, the flight SEA-ORD may be cancelled as shown in case (a) of Figure 20.13. By making this cancellation, the

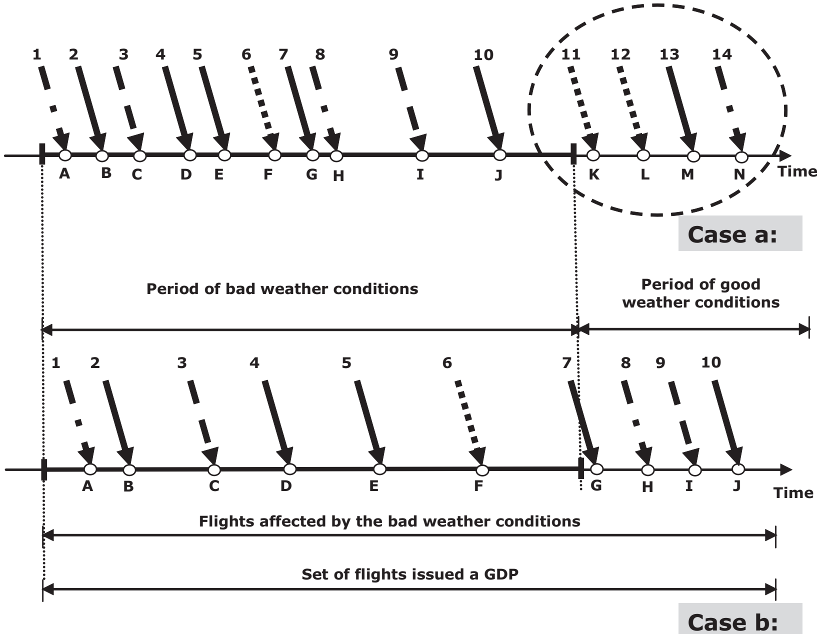


Figure 20.8 Forced flight cancellations due to GDP

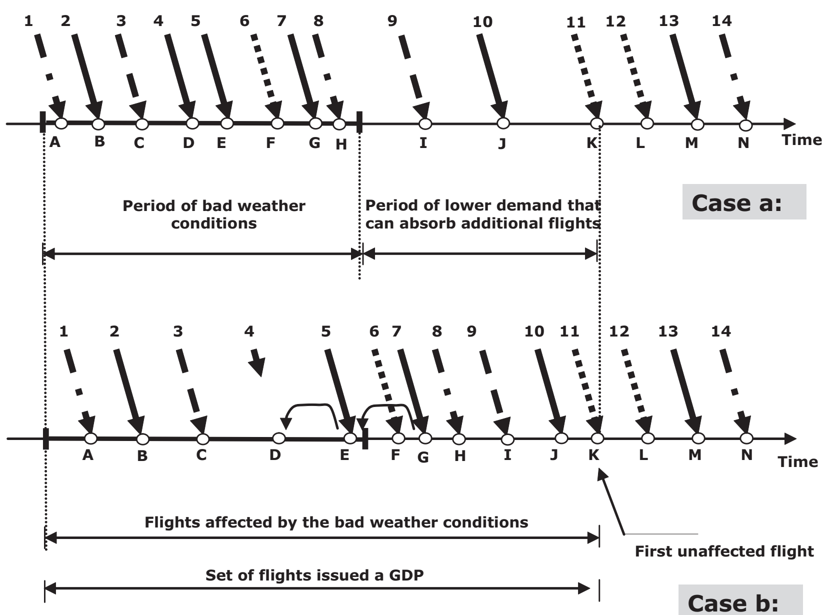


Figure 20.9 Flight cancellation as a recovery action

aircraft will stay overnight at ORD. Second, the flight ORD-SEA may be cancelled as shown in case (b). By making this cancellation, the aircraft stays overnight at SEA. Selecting which cancellation pattern to choose depends on the weather conditions overnight at both SEA and ORD. If, for example, there is an overnight snow storm at ORD, the second cancellation pattern might be preferred so the aircraft is out of harm at SEA and away from the storm at ORD.

The main problem with flight cancellation is that it misplaces the resources of the flight including the aircraft, pilots, and flight attendants at the origin station of the canceled flight. Accordingly, these resources are not able to continue on any other flights that were originally scheduled after the canceled flight. It is very expensive for the air carrier to move the empty aircraft from one station to another to reposition the aircraft. Therefore, some other flight cancellations have to occur to reposition the aircraft. To explain the repositioning of aircraft after cancellation, consider the aircraft route in Figure 20.10. If the flight departing from ORD to SEA is cancelled, the resources of the flight are stranded at ORD. The aircraft is not able to operate the next flight from SEA to ORD. If the air carrier decides to cancel the SEA-ORD flight as well, the aircraft is at ORD and ready to fly the next flight ORD-DEN, as shown in Figure 20.11.

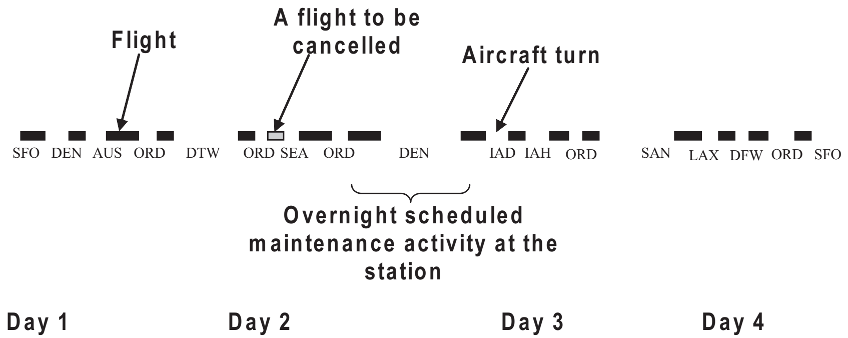


Figure 20.10 Example of flight cancellation

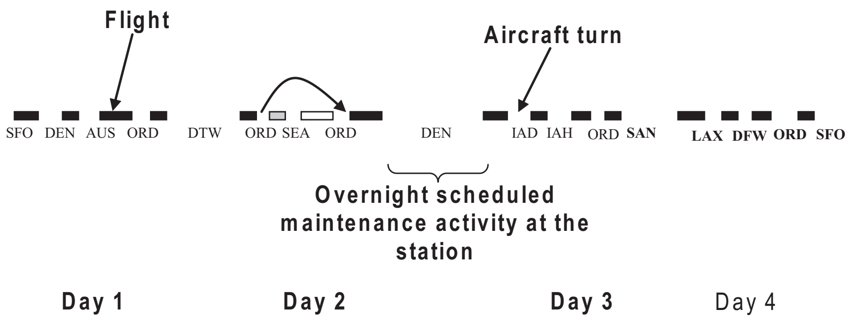


Figure 20.11 Cancellation of round trip

An air carrier might need to cancel more than one flight to reposition its aircraft after a single cancellation. For example, in Figure 20.12, when the ORD-SEA is cancelled, the air carrier might decide to also cancel the two next flights SEA-LAX and LAX-ORD. By making these cancellations, the aircraft is ready to fly the ORD-IAD flight without moving the aircraft from ORD.

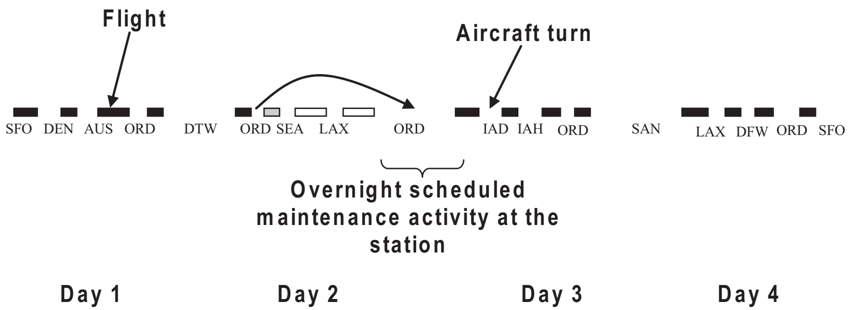


Figure 20.12 Cancellation of a loop of three flights

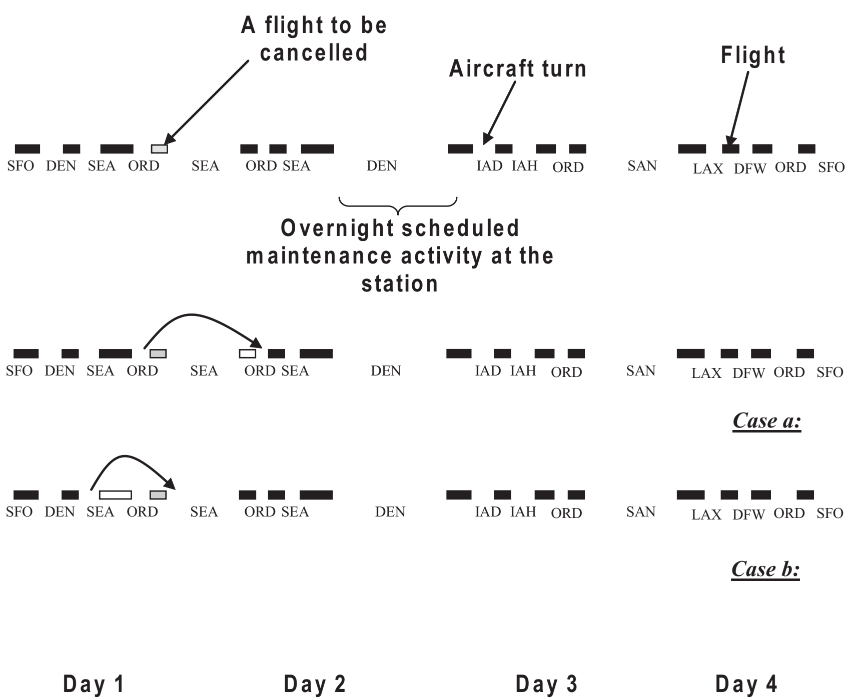


Figure 20.13 Different cancellation plans in the aircraft route

A Final Note on Irregular Operations Management

In this chapter, the main components of the irregular management problem are presented, which include resource recovery, flight delay (slot allocation), and flight cancellation. Resource recovery is required to replace the air carriers that cannot operate their scheduled flights on time or within an acceptable delay threshold. Delays are encountered when a flight cannot depart at its scheduled time because of bad weather conditions or the delay of one or more of its operating resources. A flight cancellation decision is made either when weather conditions prevent the operation of the flight or when the flight lacks at least one of its operating resources or facilities. It should be noted that these decisions cannot be implemented independently. Any action related to flight cancellation is expected to affect the resource recovery and flight delay decisions. Also, any action related to flight delay is expected to affect the flight cancellation and resource recovery. Similarly, resource recovery affects the flight delay and cancellation decisions. Figure 20.14 presents the relationship between the flight delay, cancellations, and resource recovery. As shown in the figure, there are six relationships between the three actions.

Relationship 1 indicates that flight cancellation affects the flight delay. To explain this relationship, consider the GDP given in Figure 20.9. When the air carrier that operates flights 4, 5, and 7 decides to cancel flight 4, slot D becomes open, and flight 5 can land in slot D. Similarly, when flight 5 moves to slot D, slot E becomes open, and flight 7 can land in slot E, which changes the amount of delay of flight 7. This means that the cancellation of flight 4 has an impact on

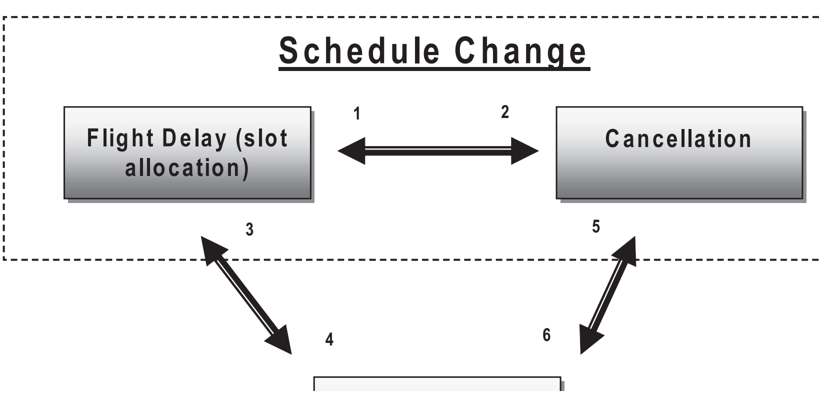


Figure 20.14 The relationship between the flight delay, cancellations and resource recovery

the delay of flights 5 and 7. Relationship 2 indicates that the flight delay affects the flight cancellation. In Figure 20.1, we have studied how the delay of flight F1 might lead to the delay of flight F4. The delay of flight F4 results in a duty limit violation for its crew. Accordingly, flight F4 has to be cancelled; otherwise, a substitute crew should be available to fly flight F4. Relationship 3 indicates that the resource recovery affects the flight delay. For example, a crew might be scheduled to stay overnight at a station before taking an early flight the next day. If the crew arrives late to this station, the next-day early flight might need to be delayed until the crew takes the minimum legal layover period. If other crews are available to handle the early flight on time—maybe a reserve or standby crew can be used—then the early flight is not delayed. If no substitute crew is available, the early flight has to be delayed until the crew completes the minimum required legal layover. Relationship 4 indicates that the flight delay affects the resource recovery. This is a straightforward relationship where a flight can be delayed to wait for its inbound delayed resource (aircraft, pilots, or flight attendants). Relationship 5 indicates that the resource recovery affects the flight cancellations. For example, the unavailability of resources might force a decision to cancel a flight. For example, in Figure 20.1, if no substitute is found for the crew of flight F4, flight F4 is cancelled because it cannot operate with a crew that has a duty limit violation. Finally, relationship 6 indicates that the flight cancellation has an impact on the resource recovery. For instance, if a flight is cancelled, its aircraft and onboard crew are stranded at the origin of the cancelled flight, and they cannot complete their scheduled downline flights. Accordingly, resource recovery decisions need to be considered at the destination of the cancelled flight to operate any downline flight for these resources. Also, the stranded resources need to be rerouted so that they can catch their original assignments or new assignments that are given to them.

Passengers Rebooking

Passengers are the main customers of the air carrier and the users of its product. All the recovery decisions of the air carrier schedule during irregular operations conditions are usually considered to alleviate the impact of the adverse weather conditions on the passengers of the different flights in the schedule. When the air carrier schedule is disrupted due to flight cancellations or severe delays, the air carrier might be able to rebook and accommodate its passengers on other flights in its schedule or on flights that belong to other air carriers. The passengers are usually ranked based on their loyalty to the air carrier (for example, passengers with high mileage in the air carrier mileage plan). Then for each passenger a new itinerary is generated, and if accepted by the passenger, the passenger is rebooked on this new itinerary, and seats are reserved in the reservation system. In some cases, the process becomes challenging when the air carrier schedule is subject to continuous changes and modifications, especially during long periods of adverse

weather conditions. Many air carriers adopt computer software that automates the passenger rebooking process.

Primary Contributions

Most work on disruption management in the airlines focuses on resolving conflicts for a single resource at a time (that is, aircraft and crew). Also, some work is reported on the integrated recovery models that focus one more than one resource. However, to our knowledge, none of these integrated models is implemented by any major airline. For an in-depth and theoretical description of the academic research within disruption management, we refer to the review by Clausen et al. (2005), Kohl et al. (2007), and Filar et al. (2001). In this section, the main contributions of schedule recovery are highlighted. First, the literature on aircraft recovery is presented followed by crew recovery models. Finally, integrated recovery models are presented.

One of the first studies for the aircraft recovery problem is the work by Teodorovic and Guberinic (1984). The objective of the model is to minimize the total passenger delays in case of aircraft unavailability. The problem is modeled based on the connection network, which consists of two types of nodes that represent the flight and the aircraft. The problem is solved by finding the shortest Hamiltonian path in the network, using a branch-and-bound algorithm. Teodorovic and Stojkovic (1990) extend the previous work to also consider the airport curfews. They present a greedy solution method that is tested on a small sample of 14 aircraft and 80 flights. Teodorovic and Stojkovic (1995) extend their model to also include the crew considerations. Their approach is based on two objectives, where the main objective is to maximize the total number of flights flown, and the second objective is to minimize the total passenger delay on flights that are not canceled. They also propose an algorithm that describes how the checks of the maintenance requirements are considered. Jarrah et al. (1993) present the solution of the aircraft recovery problem by using cancellation and re-timing. Also, the possibility of swapping aircraft is taken into consideration, where swaps can be with spare aircraft or with overnight layovers. In Mathaisel (1996), the aircraft recovery problem is modeled as a simple network flow problem. Talluri (1996) describes the problem of altering the aircraft type on a specific flight at a minimum cost. The objective is to make the change without affecting the overnight position of an aircraft mainly due to maintenance. Argüello et al. (1997) describes a heuristic approach for the reconstruction of aircraft routes when one or several aircraft are grounded. The heuristic is based on a randomized neighborhood search. Yan and Tu (1997) consider the situation of shortage of one aircraft. The problem is modeled as a timeline network, in which flights are represented by edges from origin to destination. Yan and Lin (1997) present the airline scheduling recovery problem for the temporary complete closure of airports. Similar to Jarrah et al. (1993), Cao and Kanafani (1997) make use of the timeline of the network and present a model that allows

for a solution combining delays and cancellations. The model derived is a special type of non-linear 0-1 integer program. The authors present a customized linear programming approximation algorithm for the problem. Thengvall el al. (2000) consider a model for the aircraft-schedule recovery during irregular operation with balancing user preferences. A model is developed to produce recovery schedules for single fleet recovery in case of minor disruptions; however, crew scheduling and maintenance are not taken into consideration. Also, Thengvall el al. (2001) consider a multiple fleet aircraft-schedule recovery following hub closures. Løve and Sørensen (2001) and Løve et al. (2002) investigate alternative methods for employing cancellations and re-timings in response to disruptions, which are based on local search based heuristics. Bard et al. (2001) describe a model for optimizing aircraft routings in response to groundings and delays. They present an integral minimum cost-flow model with additional constraints ensuring that a flight is either canceled or flown by a unique aircraft. Rosenberger et al. (2002) present a stochastic model for airline operations. They use a set partitioning framework that consists of a master problem and a route generating procedure. The objective is to minimize the cost of cancellation and re-timing. It is the responsibility of the controllers to define the parameters accordingly.

The literature that focuses on airline crew recovery during irregular operations is abundant. For example, Johnson et al. (1994), who represent one of the early researches in this area, present a methodology for crew rescheduling during irregular operations that is limited to solving only a single misconnection. Stojkovic (1998) formulates the crew recovery problem as an integer non-linear multi-commodity flow problem. A set partitioning formulation is used to solve the problem. The model is tested on data of one fleet from a major US carrier. Wei et al. (1997) and Song et al. (1998) develop an optimization model and an algorithm for crew management under irregular operating conditions. The problem is solved based on a space-time network, which is considered for a certain time window. Lettovsky et al. (2000) present a methodology for recovering crew in the case of disruptions. A preprocessing technique is used to extract a subset of the schedule for recovery and rescheduling. A fast crew-pairing generator is used to construct feasible crew trips. The crew recovery problem is then formulated and solved as a generalized set covering problem. Abdelghany et al. (2004) develop a decision support tool that automates the crew recovery during irregular operations for a large-scale commercial airline. The paper details and subdivides the recovery problem into four categories: misplacement problems, rest problems, duty problems, and unassigned problems. Based on detailed information regarding the current plan and pool of problems, the recovery problem is solved in steps. Several means are used for recovery, including delaying, swapping, deadheading (extra crew), and the use of standby crew. The proposed model is an assignment model with side constraints. Due to the stepwise approach, the proposed solution is suboptimal. The results are presented in a situation from a US airline with 18 problems.

One of the earliest literatures of the integrated recover models is the work of Lettovsky (1997), which presents a framework for an integrated airline recovery.

In this work, a linear mixed-integer mathematical problem that maximizes the total profit to the airline while capturing the availability of the three most important resources: aircraft, crew, and passengers is presented. The formulation has three parts corresponding to each of the resources; that is, crew assignment, aircraft routing, and passenger flow. These three parts are controlled by a master problem denoted as the schedule recovery model. Stojkovic and Soumis (2001) present a model for rescheduling aircraft and pilots simultaneously for one day. The problem is modeled as an integer non-linear multi-commodity flow model with additional constraints. The problem is solved using column generation with a master problem and a subproblem per pilot. Abdelghany et al. (2008) extend the work of Abdelghany et al. (2004) to develop an integrated recovery model that considers the three airline resources that include aircraft, pilots, and cabin crew. The model also considers flight delay as a recovery option. In this framework, flights are grouped based on their departure time in a chronological order, and several successive assignment problems are solved to assign resources to flights. After each assignment, a simulation model is used to simulate the future schedule and predict any possible resource disruptions that might result from flight delay decisions. The model is applied to the schedule of one major airline in the US Several experimental scenarios are performed, and the running time of the model under the different scenarios is presented.

References

Abdelghany, A., Ekollu, G., Narasimhan, R., and Abdelghany, K. 2004. A Proactive Crew Recovery Decision Support Tool for Commercial Airlines during Irregular Operations. *Annals of Operations Research*, 127, 309-331.

Abdelghany, K., Abdelghany, A., and Ekollu, G. 2008. An Integrated Decision Support Tool for Airlines Schedule Recovery during Irregular Operations. *European Journal of Operational Research*, 185 (2), 825-848.

Argüello, M., Bard, J., and Yu, G. 1997. A GRASP for Aircraft Routing in Response to Groundings and Delays. *Journal of Combinatorial Optimization*, 5, 211-228.

Bard, J., Yu, G., and Argüello, M. 2001. Optimizing Aircraft Routings in Response to Groudings and Delays. *IIE Transactions*, 33, 931-947.

Cao, J. and Kanafani, A. 1997. Real-Time Decision Support for Integration of Airline Flight Cancellations and Delays Part I. *Transportation Planning and Technology*,20, 183-199.

Cao, J. and Kanafani, A. 1997. Real-Time Decision Support for Integration of Airline Flight Cancellations and Delays Part II. *Transportation Planning and Technology*, 20, 201-217.

Clausen, J., Larsen, A., and Larsen, J. 2005. *Disruption Management in the Airline Industry—Concepts, Models and Methods*. In series: IMM-Technical Report-2005-01, Technical University of Denmark, Denmark.

Filar, J.A., Manyem, P., and White, K. 2001. How Airlines and Airports Recover from Schedule Perturbations: A Survey. *Annals of Operations Research*, 108(1-4), 315-333.

Jarrah, A., Yu, G. Krishnamurthy, N., and Rakshit, A. 1993. A Decision Support Framework for Airline Flight Cancellations and Delays. *Transportation Science*, 27, 266-280.

Johnson, E., Lettovsky, L., Nemhauser, G., Pandit, R., and Querido, S. 1994, *Final Report to Northwest Airlines on the Crew Recovery Problem*, Technical Report, The Logistic Institute, Georgia Institute of Technology, Atlanta, GA.

Kohl, N.A., Larsen, J., Larsen, A., Ross, and Tiourine S. 2007, Airline disruption management—Perspectives, experiences and outlook. *Journal of Air Transport Management*, 13(3), 149-162.

Lettovsky, L. 1997. *Airline Operations Recovery: An Optimization Approach*. PhD thesis, Georgia Institute of Technology, Atlanta, USA.

Lettovsky, L., Johnson, E.L., and Nemhauser, G.L. 2000. Airline Crew Recovery. *Transportation Science*, 34(4), 337-348.

Løve, M. and Sørensen, K. R. 2001, *Disruption management in the airline industry*, Master's thesis, *Informatics and Mathematical Modelling* (IMM). Techical University of Denmark (DTU), accessed March 2001 from http://www.imm.dtu.dk/documents/ftp/ep2001/ep16_01-a.html.

Løve, M., Sørensen, K. R., Larsen, J., and Clausen, J. 2002, Disruption Management for an Airline — Rescheduling of Aircraft, In Stefano Cagnoni, Jens Gottlieb, Emma Hart, Martin Middenhof, and Gunther R. Raidl, (Eds), *Applications of Evolutionary Computing*, volume 2279 of *Lecture Notes in Computer Science*, Springer, pages 315-324.

Mathaisel, D.F.X. 1996. Decision Support for Airline System Operations Control and Irregular Operations. *Computers and Operations Research* 23, 1083-1098.

Rosenberger, J., Schaefer, A., Goldsmans, D., Johnson, E., Kleywegt, A., and Nemhauser, G. 2002. A stochastic model of airline operations. *Transportation Science*, 36(4), 357-377.

Song, M., Wei G., and Yu G., 1998, A Decision Support Framework for Crew Management During Airline Irregular Operations, In Gang Yu, (ed.), *Operations Research in the Airline Industry,* Kluwer Academic Publishers, Boston.

Stojkovic, M. and Soumis, F. 2001. An Optimization Model for the Simultaneous Operational Flight and Pilot Scheduling Problem. *Management Science*, 47(9), 1290- 1305.

Stojkovic, M., Soumis, F., and Desrosiers, J. 1998. The Operational Airline Crew Scheduling Problem. *Transportation Science*, 32, 232-245.

Talluri, K.T. 1996. Swapping Applications in a Daily Airline Fleet Assignment. *Transportation Science*, 30, 237-248.

Teodorovic, D. and Guberinic, S. 1984. Optimal Dispatching Strategy on an Airline Network after a Schedule Perturbation. *European Journal of Operational Research*, 15, 178-182.

Teodorovic, D. and Stojkovic, G. 1990. Model for Operational Daily Airline Scheduling. *Transportation Planning and Technology*, 14, 273-285.

Teodorovic, D. and Stojkovic, G. 1995. Model to Reduce Airline Schedule Disturbances. *Journal of Transportation Engineering*, 121, 324-331.

Thengvall, B.G., Bard, J., and Yu, G. 2000. Balancing User Preferences for Aircraft Schedule Recovery during Irregular Operations. *IIE Transactions*, 32, 181-193.

Thengvall, B.G., Yu, G., and Bard, J.F. 2001. Multiple Fleet Aircraft Schedule Recovery Following Hub Closures. *Transportation Research Part A*, 35, 289-308.

Wei, G., Yu, G. and Song, M. 1997. Optimization Model and Algorithm for Crew Management during airline irregular operations. *Journal of Combinatorial Optimization* 1, 305–321.

Yan, S. and Lin, C. 1997. Airline Scheduling for the Temporary Closure of Airports. *Transportation Science*, 31, 72-82.

Yan, S. and Tu, Y. 1997. Multifleet Routing and Multistop Flight Scheduling for Schedule Perturbation. *European Journal of Operational*, 103(1), 155-169.

This page has been left blank intentionally

Index

A
activity selection problem, 113
agreement, 42, 213, 221–224
air traffic control, 17, 129
Aircraft cleaning, 4, 7, 59, 79, 244
aircraft rotation, 81–82, 87
aircraft routing, 4, 8–9, 13, 79, 81–87, 92, 94, 121, 271
aircraft swapping, 87, 254–255, 257
aircraft turn, 13–14, 79, 80, 112–114, 244, 265
airport capacity, 229
airway, 129, 130–131
alliance, 3–4, 41, 221, 225–226
altitude, 104, 107, 129, 130–134
apron, 109, 112–114
authorization level, 150, 173–175, 184

B
baggage handling, 3, 7, 15, 17, 109, 119–122, 124–125, 127–128, 250
baggage transfer, 119, 124, 126, 127
baggage-claim area, 125
bidlines, 199, 104, 107
block hours, 80, 85, 100–101
booking classes, 147–148, 150, 157, 165, 171, 181, 184–186, 189, 190–191, 193, 202, 221
booking fees, 207–209
booking request, 145–147, 153, 182, 189–190, 194
briefing period, 91
business travelers, 15–16, 17, 22, 31, 35–36, 142–143, 147, 153, 159, 181, 213–214

C
cabin crew, 4, 7, 9, 243, 271
cancellation, 110, 142, 147, 156, 173, 178, 232, 239, 243, 250–251, 253–254, 258, 263, 265–268
cargo, 1–2, 7, 14, 59, 63, 74, 79, 109, 111, 129, 131, 243, 246
cargo handling, 109, 111
catering, 4, 7, 14, 79, 244
city baggage, 119, 124
Cluster analysis, 175
code share, 28–29, 33, 39, 41, 45, 150, 221
code-share agreements, 39, 42, 221–225
collaborative decision making, 229, 233–234, 259
column generation, 85, 87, 94, 98, 102–104, 271
commission, 205
computer reservation systems, 206
congestion, 3, 10, 135, 237
connecting passengers 10, 39, 109–110, 112, 114, 119, 152, 189, 250–251, 263
connection time, 5, 14, 28–29, 33, 42, 45, 89, 109, 244–245, 257–258
continuity constraints, 65
contracts, 213–215, 218–219, 258
corporation, 213–215, 219, 241
coverage constraint, 63
crew cost, 15, 63, 89
crew legalities, 92
crew pairing, 89–93, 101–102, 104, 121, 270
crew planning, 14, 74, 89, 102
crew rostering, 89–90, 99–101, 104–105
crew scheduling, 8, 17, 74, 89, 98, 102–103, 270
crew seniority, 99
crew workload, 89
customer service, 3–4, 7, 15, 243, 245, 250

D
days away from home, 101
deadhead, 98–99, 101, 257
delay, 5, 8, 12, 17, 35, 79, 110, 127, 130, 229–230, 232, 235, 237, 239, 243, 245–251, 253–254, 256–263, 267–269, 271
demand data, 41, 43
Demand forecasting, 5, 12, 16, 40, 145, 146
Demand modeling, 17, 34, 49
deregulation, 141, 206
discount, 186, 213–214, 219, 223
Discrete Choice Theory, 24–25
dispatcher, 7, 129, 132
displacement cost, 203, 214–216, 218–219, 223–224
distribution channels, 16, 40, 49, 205, 208–210
dual price, 193–194
duty period, 14, 89–91, 98, 244–245, 250, 256–257
dynamic programming, 134, 137, 178–179, 187, 202–203, 237

E
elastic demand, 142
expected marginal revenue, 216
expected revenue, 146, 168, 182, 184, 186, 191, 216
expected seat revenue, 168, 170, 182, 184–185, 216, 218

F
fare classes, 146–148, 150, 156–157, 162, 165, 179, 181, 184–186, 190–195, 200–203
Federal Aviation Administration, 80, 89, 229, 234, 259
fleet assignment, 8, 13, 17, 49, 53–55, 61, 63–66, 71–72, 74–76, 79, 81
fleet characteristics, 53
flight capacity, 143, 185, 193, 198, 200, 216
flight path, 129–131
flight planning, 17, 129–133, 135, 137
flying hours, 14, 71
free sale, 221
fuel, 3, 5, 7, 13, 17, 53, 63, 89, 129, 130–132, 134–138
fuel consumption, 13, 53, 129, 130–131, 134
fuel ferrying, 135–136, 138

G
gate assignment, 4, 109–111, 113–115, 116, 121
gate walkways, 109
global distribution systems, 16, 206, 208
global optimum, 133
greedy approach, 113
ground delay program, 229–230
ground holding, 237–239, 245
ground time, 59, 79, 82, 91, 243

H
hard block, 221
Herfindahl-Hirschman Index, 47
historical data, 21, 26, 147, 170
hub-and-spoke, 8–12, 86, 92, 94, 150–152, 189–190, 252
hub-spoke network, 262

I
identification tag, 119
idle time, 14, 79, 110–111, 122–124, 247
incremental revenue, 193–194, 214–215, 222, 225
interconnection node, 60–61, 65, 67, 70
Internet, 205, 208, 210
itinerary builder, 42–43
itinerary-fare class, 157–160, 162, 165–166, 168, 170, 190–195, 198, 200–202

L
landing slots, 3, 229, 231–233, 235–236, 239, 249, 263
layover, 14, 22, 89, 90–91, 244–245, 257 268
leisure travelers, 15–16, 22, 142–143, 147, 152, 159, 181
level of service, 27, 36, 115
line of flying, 15, 79
linear program, 74, 104, 137, 192–193, 203
load factor, 141, 143, 147, 221

low cost carriers, 141
loyalty, 22, 209, 268

M
maintenance activity, 13–14, 68–69, 80, 85, 243–244, 255, 265
maintenance constraint, 68
market penetration, 49, 205, 210
Market share, 43, 47
marketing carrier, 221, 223–224
minimum legal rest, 245

N
negotiation, 4, 213–214, 222
nesting, 73, 147–150, 184, 191, 193–194, 202
network coverage, 3, 9, 40, 221
network structure, 8–12, 48, 85, 92, 94, 132, 150, 181, 184, 189, 221, 262
no-show rate, 173–175

O
Official Airline Guide, 41
operating carrier, 221–224
operation management center, 259
or line maintenance, 80
Overbooking, 153, 173–174, 176–177, 179, 187
overflight cost, 130

P
pairing, 89–93, 96, 101–102, 104, 121, 270
parallel nesting, 150
passenger re-booking, 17, 269
piers, 119–124, 128
point-to-point, 8, 12, 150, 152, 181, 184
proactive approach, 253
probability distribution, 166–168, 176–178, 182
Protected Seats, 148, 153, 184–185
protection level, 147–148, 202

Q
quads, 119

R
ramp agents, 4, 7, 243
random variable, 116
ration-by-schedule, 235
reclear, 135
recovery desk, 259
redispatch, 135
regulations, 3, 16–17, 89–91, 100, 129, 229, 254, 257
Reserve crew, 14, 90, 256
revenue management, 9, 15, 49, 141, 150, 152, 155, 171, 179, 181, 189

S
safety regulations, 129, 229
schedule compression, 235
schedule data, 41–42, 233
schedule disruptions, 15, 254
seat inventory control, 4, 8, 16–17, 145–146, 153, 158, 165, 181, 186–187, 189–191, 195, 221
sequential nesting, 147, 148–150
Sequential Recovery, 258
set partitioning, 74, 83–86, 92–96, 98–99, 102–104, 270
simulator training, 101
sizing constraint, 64
slack time, 246–250, 253, 260
slot allocation, 239, 260, 262, 267
snapshot, 157–159
socioeconomic characteristics, 22, 35
soft block, 221
sorting facility, 119–120, 125–126
standby crew, 90, 256, 268, 270
stranded crew, 257
substitution, 36, 235–237
swapping, 16, 87, 254,–257, 269–270

T
Tabu-search, 113–114, 116
taxi distance, 109
taxi time, 111
taxi-in, 10, 251
taxi-out, 10, 251
through flight, 66–67, 74
through traffic, 79, 82, 85
ticket distribution, 16–17, 39–40, 49, 205, 208–210
time bank, 10, 110
time-series analysis, 160–162
time-staged flight network, 55, 59–60
timetable, 53–55, 63, 75

time-window, 69–71
training, 89–90, 99–101
transfer baggage, 119, 124–127
travel agents, 16, 205–209
trip characteristics, 22, 25
trippair, 14–15, 89–96, 244–245, 257
turn time, 13, 79–80, 85, 91, 109, 110
unconstrained demand, 165

U
utility function, 25–26, 32

W
walking distance, 109, 111–115
waypoints, 129–130, 133, 137
weather, 5, 8, 15–16, 129–132, 160, 174, 229, 231–232, 247, 254, 259, 263, 265, 267–269
wholesalers, 205
willingness-to-pay, 143
wind, 129, 131–134
work rules, 16, 89, 91, 94, 244, 254, 258